PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12P 21/06, 21/04, C12N 15/00, 15/09, 15/63, 15/70, 15/74, C07H 21/02, 21/04 — A1

(11) International Publication Number: WO 97/31123

(43) International Publication Date: 28 August 1997 (28.08.97)

(21) International Application Number: PCT/US97/02952

(22) International Filing Date: 20 February 1997 (20.02.97)

(30) Priority Data:
60/012,028 21 February 1996 (21.02.96) US

(60) Parent Application or Grant
(63) Related by Continuation
US 60/012,028 (CIP)
Filed on 21 February 1996 (21.02.96)

(71) Applicant *(for all designated States except US)*: BOARD OF REGENTS, THE UNIVERSITY OF TEXAS SYSTEM [US/US]; 201 West 7th Street, Austin, TX 78701 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: NORRIS, Steven, J. [US/US]; 5314 Stillbrooke Drive, Houston, TX 77096 (US). ZHANG, Jing-Ren [CN/US]; 9220 Bellwood, Houston, TX 77036 (US). HARDHAM, John, M. [US/US]; 5902 DeMoss Drive, Houston, TX 77081 (US). HOWELL, Jerrilyn, K. [US/US]; 6837 Langdon Lane, Houston, TX 77074 (US). BARBOUR, Alan, G. [US/US]; 34 Whitman Court, Irvine, CA 92612 (US). WEINSTOCK, George, M. [US/US]; 5046 Wigton, Houston, TX 77096 (US).

(74) Agent: KITCHELL, Barbara, S.; Arnold, White & Durkee, P.O. Box 4433, Houston, TX 77210 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: VMP-LIKE SEQUENCES OF PATHOGENIC BORRELIA

(57) Abstract

The present invention relates to DNA sequences encoding Vmp-like polypeptides of pathogenic *Borrelia*, the use of the DNA sequences in recombinant vectors to express polypeptides, the encoded amino acid sequences, application of the DNA and amino acid sequences to the production of polypeptides as antigens for immunoprophylaxis, immunotherapy, and immunodiagnosis. Also disclosed are the use of the nucleic acid sequences as probes or primers for the detection of organisms causing Lyme disease, relapsing fever, or related disorders, and kits designed to facilitate methods of using the described polypeptides, DNA segments and antibodies.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgyzstan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic			SE	Sweden
CG	Congo	KR	Republic of Korea	SG	Singapore
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LR	Liberia	SZ	Swaziland
CS	Czechoslovakia	LT	Lithuania	TD	Chad
CZ	Czech Republic	LU	Luxembourg	TG	Togo
DE	Germany	LV	Latvia	TJ	Tajikistan
DK	Denmark	MC	Monaco	TT	Trinidad and Tobago
EE	Estonia	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	UG	Uganda
FI	Finland	ML	Mali	US	United States of America
FR	France	MN	Mongolia	UZ	Uzbekistan
GA	Gabon	MR	Mauritania	VN	Viet Nam

VMP-LIKE SEQUENCES OF PATHOGENIC *BORRELIA*

1.0 BACKGROUND OF THE INVENTION

1.1 Field of the Invention

The invention relates to the field of molecular biology; in particular, to immunogenic compositions and recombinant VMP-like genes useful for treatment and diagnosis of Lyme disease. Also included are methods for the determination of virulence factors in Lyme disease.

1.2 Description of Related Art

Lyme disease is a bacterial infection caused by pathogenic spirochetes of the genus *Borrelia*. The infection can occur in humans, dogs, deer, mice and other animals, and is transmitted by arthropod vectors, most notably ticks of the genus *Ixodes*. *Borrelia burgdorferi*, the most common cause of Lyme disease in North America, was first cultured in 1982. *B. garinii* and *B. afzelii* are the most common infectious agents of Lyme disease in Europe, and another species, *B. japonicum*, has been described in Japan. These organisms are closely related and cause similar manifestations with multiple stages: an expanding rash at the site of the tick bite (erythema migrans); fever, lymphadenopathy, fatigue, and malaise; effects of disseminated infection, including carditis, meningoradiculitis, and polyarthritis; and chronic manifestations including arthritis and neurologic disorders. Lyme disease is often difficult to diagnose because of shared manifestations with other disorders, and it can also be refractory to treatment during late stages of the disease. It is most common in areas such as suburban regions of upstate New York and Connecticut, where large populations of deer and white-footed mice serve as the principal mammalian hosts and reservoirs of infection. Approximately 10,000 cases of Lyme disease in humans are reported per year in the United States, and it is also a significant veterinary problem due to a high infection rate of dogs and other domestic animals in endemic regions.

B. burgdorferi, the etiologic agent of Lyme disease, is able to persist for years in patients or animals despite the presence of an active immune response (Steer, 1989; Schutzer, 1992). Antigenic variation has been postulated previously as a mechanism whereby *B. burgdorferi* evades the immune response in the mammalian host (Schwan *et al.*, 1991; Wilske *et al.*, 1992). Antigenic variation has been defined as changes in the structure or expression of antigenic proteins that occurs during infection at a frequency greater than the usual mutation rate (Borst and Geaves, 1987; Robertson and Meyer, 1992; Seifert and So, 1988).

Relapsing fever is another disease caused by pathogenic *Borrelia*. It has both epidemic and endemic forms. The epidemic form is caused by *B. recurrentis* and is transmitted between humans by lice. It was a major source of morbidity and mortality during World War I, but has been rare since then due largely to public health measures. Endemic relapsing fever is an epizootic infection caused by several *Borreliae* species, including *B. hermsii*. It occurs sporadically among hunters, spelunkers, and others who come in contact with infected soft-bodied ticks of the genus *Ornithidorus*. Relapsing fever is characterized by two or more episodes or "relapses" of high bacteremia (up to 10^8/ml). The first wave of infection is caused by *Borreliae* expressing a certain Variable Major Protein (VMP) on their surface (*e.g.* Vmp21). The gene encoding this VMP is located at a promoter site in the expression plasmid, whereas over 24 nonexpressed copies of different VMP genes are present on the so-called silent plasmid. When the host develops antibodies against the expressed VMP, the organisms of that stereotype are destroyed and the patient improves. However, a small proportion of organisms have undergone antigenic switching to a different stereotype. Nonreciprocal recombination occurs between the expression plasmid and the silent plasmid, resulting in the insertion of a different VMP gene in the expression site (*e.g.*, Vmp7). The organisms expressing Vmp7 are not affected by the anti-Vmp21 antibodies, and therefore multiply in the host and cause a second episode of the disease. Up to five of these 3-5 day episodes can occur, separated by 1-2 week intervals.

Such well-demarcated episodes of infection do not occur during Lyme disease, and fewer organisms are present in the blood and in tissues at any stage. However, there are reasons to suspect that similar mechanisms of antigenic variation may occur in *B. burgdorferi* and other

Lyme disease *Borreliae*. The infection, if untreated, commonly persists for months to years despite the occurrence of host antibody and cellular responses; this observation indicates effective evasion of the immune response. Lyme disease may be disabling (particularly in its chronic form), and thus there is a need for effective therapeutic and prophylactic treatment.

Certain *B. burgdorferi* genes and proteins have been patented, including Outer Surface Protein D (OspD) (U.S. Patent No. 5,246,844; issued September 21, 1993). OspD has not proven to be a useful protein for diagnosis or immunoprotection. Other proteins, including OspA and OspC, have been considered as vaccine candidates for Lyme disease, including a recombinant OspA vaccine currently in human clinical trials. Other vaccines are in use or undergoing testing in veterinary applications, including vaccination of dogs. However, animal studies indicate that OspA vaccination may not be effective against all strains of Lyme disease *Borreliae*. OspA is also not useful for immunodiagnosis, due to weak antibody responses to OspA in Lyme disease patients.

Previous studies have generally failed to provide evidence for the occurrence of antigenic variation in Lyme disease *Borrelia*. Genetic heterogeneity in the genes encoding the membrane lipoproteins OspA, OspB, OspC, and OspD has been well documented among strains of Lyme disease *Borreliae* (Marconi *et al.*, 1993; Marconi *et al.*, 1994; Livey *et al.*, 1995). In addition, mutations in *ospA* and *ospB* have been shown to occur *in vitro* (Rosa *et al.*, 1992; Sadziene *et al.*, 1992). However, no significant antigenic change (Barthold, 1993) or gross genetic alteration (Persing *et al.*, 1994; Stevenson *et al.*, 1994) has been detected in *B. burgdorferi* N40 isolates from chronically infected BALB/c and C3H mice, other than the loss of the 38-kilobase (kb) plasmid encoding OspD. Therefore the heterogeneity in Osp proteins observed among *B. burgdorferi* sensu lato isolates appears to represent evolutionary divergence ("antigenic drift") rather than antigenic variation.

There is a commercial demand for vaccines and diagnostic kits for Lyme disease, both for human and veterinary use. Several companies have active research and development programs in these areas.

2.0 SUMMARY OF THE INVENTION

Partial and complete DNA sequences have been determined for several recombinant clones containing DNA encoding VMP-like sequences. The identification and characterization of these sequences now allows: (1) identification of the expressed gene(s) in *B. burgdorferi*; (2) expression of these gene(s) by a recombinant vector in a host organism such as *E. coli*; (3) immunization of laboratory animals with the resulting polypeptide, and determination of protective activity against *B. burgdorferi* infection; (4) use of antibodies against the expressed protein to identify the reactive polypeptide(s) in *B. burgdorferi* cells; (5) use of the expressed protein(s) to detect antibody responses in infected humans and animals; (6) determination of the presence, sequence differences, and expression of the VMP-like DNA sequences in other Lyme disease *Borreliae*.

The invention is contemplated to be useful in the immunoprophylaxis, diagnosis, or treatment of Lyme disease, relapsing fever, or related diseases in humans or animals. It is expected that recombinant or native proteins expressed by the VMP-like genes (or portions thereof) will be useful for (a) immunoprophylaxis against Lyme disease, relapsing fever, or related disorders in humans and animals; (b) immunotherapy of existing Lyme disease, relapsing fever, or related illnesses, by way of immunization of injection of antibodies directed against VMP-like proteins; and (c) immunodiagnosis of Lyme disease, relapsing fever, or related diseases, including their use in kits in which the VMP-like proteins are the sole antigen or one of multiple antigens. The DNA may be employed in: (a) production of recombinant DNA plasmids or other vectors capable of expressing recombinant polypeptides; and (b) design and implementation of nucleic acid probes or oligonucleotides for detection and/or amplification of VMP-like sequences. The latter is expected to have application in the diagnosis of infection with *Borrelial* organisms.

Similar sequences in *B. burgdorferi* and other Lyme disease *Borreliae* have not been reported previously, as determined by BLAST searches of current nucleotide and amino acid databases including Genbank, the EMBL DNA database, and the Swiss Protein database. Although there is some similarity between the *B. burgdorferi* deduced amino acid sequences with previously published *B. hermsii* VMP deduced amino acid sequences, the degree of identity and similarity is only ~30% and ~50%, respectively. Outer surface protein C (OspC) of

Lyme disease organisms has been reported to have sequence similarities to VMPs, but the highest similarity is to a different subgroup of VMPs than the sequences reported here (Carter *et al.*, 1994). The VMP-like sequences such as those contained in pJRZ53-31 have a low degree of homology with OspC from some Lyme disease organisms (*e.g. B. burgdorferi* 2591), as indicated by a BLASTP homology score of 60 and a probability of 0.0013. Thus, the *B. burgdorferi* VMP-like DNA sequences are unique, although they have an apparent evolutionary relationship with other *Borrelia* genes.

Another aspect of the invention is the method for identification of possible virulence factors. This approach entails subtractive hybridization of target DNA from high infectivity organisms with driver DNA from low-infectivity strains or clones. This procedure greatly enriches for sequences which differ between the high- and low-infectivity strains and thus may encode proteins important in virulence. Of particular utility is the use of closely related isogenic clones that differ in their infectivity; in this case, the DNA differences should be restricted more stringently to those related to infectivity.

Open reading frames in a *B. burgdorferi* plasmid that encode hypothetical proteins resembling the VMP proteins of relapsing fever organisms have now been identified. The inventors have found that the presence of the plasmid containing these VMP-like sequences in *B. burgdorferi* clones correlates strongly with infectivity. Thus it is likely that the proteins encoded by the VMP-like sequences are important in immunoprotection and pathogenesis. Proteins encoded by the VMP-like sequences of *B. burgdorferi* may provide protection when used either alone or in combination with other antigens. They may also be useful for immunodiagnosis.

The invention is considered to include DNA segments corresponding to 20, 30, and 40 base pairs of the VMP-like sequences; DNA segments inclusive of the entire open reading frames of the VMP-like sequences; amino acid sequences corresponding to both conserved and variable regions of the VMP-like sequences; recombinant vectors encoding an antigenic protein corresponding to the above amino acid sequences; recombinant cells where extrachromosomal DNA expresses a polypeptide encoded by the DNA encoding *Borrelia* VMP-like sequences; a recombinant *B. burgdorferi* or *E. coli* cell containing the DNA encoding VMP-like sequences; methods of preparing transformed bacterial host cells using the DNA encoding the VMP-like

polypeptides; methods using the plasmid or vector to transform the bacterial host cell to express *B. burgdorferi* polypeptides encoded by the DNA sequences; methods for immunization of humans or animals with the native *B. burgdorferi* polypeptide or polypeptides expressed by recombinant cells that include DNA encoding the VMP-like polypeptides; and methods for identifying potential virulence factors using subtractive hybridization between target DNA from high-infectivity cells and driver DNA from low-infectivity cells.

Also included in the invention are primer sets capable of priming amplification of the VMP-like DNA sequences; kits for the detection of *B. burgdorferi* nucleic acids in a sample, the kits containing a nucleic acid probe specific for the VMP-like sequences, together with a means for detecting a specific hybridization with the probe; kits for detection of antibodies against the VMP-like sequences of *B. burgdorferi* and kits containing a native or recombinant VMP-like polypeptide, together with means for detecting a specific binding of antibodies to the antigen.

2.1 Methods of Treatment

An important aspect of the invention is the recognition that *Borrelia* VMP-like sequences recombine at the *vls* site, with the result that antigenic variation is virtually limitless. Multiclonal populations therefore can exist in an infected patient so that immunological defenses are severely tested if not totally overwhelmed. Thus there is now the opportunity to develop more effective combinations of immunogens for protection against *Borrelia* infections or as preventive inoculations such as in the form of cocktails of multiple antigenic variants based on a base series of combinatorial VMP-like antigens.

VMP-like protein preparations may be administered in several ways, either locally or systematically in pharmaceutically acceptable formulations. Amounts appropriate for administration are determined on an individual basis depending on such factors as age and sex of the subject, as well as physical condition and weight. Such determinations are well within the skill of the practitioner in the medical field.

Other methods of administration may include injection of *Borrelia* VMP-like DNAs into vaccine recipients (human or animal) driven by an appropriate promoter such as CMV, (so called DNA vaccines). Such preparations could be injected directly into lesions or transplanted

into patients for systemic immunization. DNA vaccinations techniques are currently well past the initial development stage and have shown promise as vaccination strategies.

2.2 VMP-like Genes

Recombinant proteins and polypeptides encoded by isolated DNA segments and genes are often referred to with the prefix "r" for recombinant. As such, DNA segments encoding rVMPs, or rVMP-related genes, *etc.* are contemplated to be particularly useful in connection with this invention. Any recombinant *vls* combining any of the *vls*E expression site loci and/or silent *vls* cassettes (*vls2-vls-16)* gene would likewise be very useful with the methods of the invention.

Isolation of the DNA encoding VMP polypeptides allows one to use methods well known to those of skill in the art and as herein described to make changes in the codons for specific amino acids such that the codons are "preferred usage" codons for a given species. Thus for example, preferred codons will vary significantly for bacterial species as compared with mammalian species; however, there are preferences even among related species. Shown below is a preferred codon usage table human. Isolation of spirochete DNA encoding VMP will allow substitutions for preferred human codons, although expressed polypeptide product from human DNA is expected to be homologous to bacterial VMP and so would be expected to be structurally and functionally equivalent to VMP isolated from a spirochete. However, substitutions of preferred human codons may improve expression in the human host, thereby improving the efficiency of potential DNA vaccines.

TABLE 1

Homo sapiens

Codon	υ[b]	Total #[a]	Codon	υ[b]	Total #[a]	Codon	υ[b]	Total #[a]	Codon	υ[b]	Total #[a]
UUU	16.6	72711	UCU	14.0	62953	UAU	12.3	55039	UGU	9.5	42692
UUC	21.4	95962	UCC	17.7	79482	UAC	17.0	76480	UGC	12.8	57368
UUA	6.3	28202	UCA	10.7	48225	UAA	0.7	2955	UGA	1.2	5481
UUG	11.5	51496	UCG	4.4	19640	UAG	0.5	2181	UGG	13.5	59982
CUU	11.7	52401	CCU	16.7	74975	CAU	9.6	43193	CGU	4.6	20792
CUC	19.5	87696	CCC	20.0	89974	CAC	14.6	65533	CGC	11.0	49507
CUA	6.3	28474	CCA	16.2	72711	CAA	11.4	51146	CGA	5.9	26551
CUG	40.6	182139	CCG	6.9	30863	CAG	33.7	151070	CGG	11.3	50682
AUU	15.7	70652	ACU	12.8	57288	AAU	16.6	74401	AGU	11.1	49894
AUC	23.7	106390	ACC	21.1	94793	AAC	21.1	94725	AGC	19.1	85754
AUA	6.7	30139	ACA	14.7	66136	AAA	23.2	104221	AGA	10.8	48369
AUG	22.6	101326	ACG	6.7	30059	AAG	33.9	152179	AGG	10.9	48862
GUU	10.6	47805	GCU	18.7	83800	GAU	22.0	98712	GCU	11.2	50125
GUC	15.6	70189	GCC	29.2	130966	GAC	27.0	121005	GGC	24.0	107571
GUA	6.6	29659	GCA	15.3	68653	GAA	27.8	124852	GGA	16.9	75708
GUG	30.0	134750	GCG	7.5	33759	GAG	40.8	182943	GGG	16.7	74859

Coding GC 52.96% 1st letter GC 55.98% 2nd letter GC 42.29% 3rd letter GC 60.60%

[a] Total 4489120

[b] υ = Frequency per 1000

The definition of a "VMP-like gene", "VMP-related gene" as used herein, is a gene that hybridizes, under relatively stringent hybridization conditions (see, *e.g.*, Maniatis *et al.*, 1982), to DNA sequences presently known to include related gene sequences.

To prepare an VMP-like gene segment or cDNA one may follow the teachings disclosed herein and also the teachings of any of patents or scientific documents specifically referenced herein. One may obtain a rVMP- or other related-encoding DNA segments using molecular biological techniques, such as polymerase chain reaction (PCR™) or screening of a cDNA or genomic library, using primers or probes with sequences based on the above nucleotide sequence. Such fragments may be readily prepared by, for example, directly synthesizing the fragment by chemical means, by application of nucleic acid reproduction technology, such as the PCR™ technology of U.S. Patents 4,683,195 and 4,683,202 (herein incorporated by reference). The practice of these techniques is a routine matter for those of skill in the art, as taught in various scientific texts (see *e.g.*, Sambrook *et al.*, 1989), incorporated herein by reference. Certain documents further particularly describe suitable mammalian expression vectors, *e.g.*, U.S. Patent 5,168,050, incorporated herein by reference. The VMP genes and DNA segments that are particularly preferred for use in certain aspects of the present methods are those encoding VMP and VMP-related polypeptides.

It is also contemplated that one may clone other additional genes or cDNAs that encode a VMP or VMP-related peptide, protein or polypeptide. The techniques for cloning DNA molecules, *i.e.*, obtaining a specific coding sequence from a DNA library that is distinct from other portions of DNA, are well known in the art. This can be achieved by, for example, screening an appropriate DNA library which relates to the cloning of a *vls* gene such as from the variable region of that gene. The screening procedure may be based on the hybridization of oligonucleotide probes, designed from a consideration of portions of the amino acid sequence of known DNA sequences encoding related *Borrelia* proteins. The operation of such screening protocols is well known to those of skill in the art and are described in detail in the scientific literature, for example, see Sambrook *et al.*, 1989.

Techniques for introducing changes in nucleotide sequences that are designed to alter the functional properties of the encoded proteins or polypeptides are well known in the art, *e.g.*, U.S. Patent 4,518,584, incorporated herein by reference, which techniques are also described in

further detail herein. Such modifications include the deletion, insertion or substitution of bases, and thus, changes in the amino acid sequence. Changes may be made to increase the VMP activity of a protein, to increase its biological stability or half-life, to change its glycosylation pattern, and the like. All such modifications to the nucleotide sequences are encompassed by this invention.

2. 3 VMP-Encoding DNA Segments

The present invention, in a general and overall sense, also concerns the isolation and characterization of novel *vls* gene segments, which encode combinatorial mosaics of expressed and silent regions of the *vls* gene. A preferred embodiment of the present invention is a purified nucleic acid segment that encodes a protein that has at least a partial amino acid sequence in accordance with SEQ ID NO:2. Another embodiment of the present invention is a purified nucleic acid segment, further defined as including nucleotide sequences in accordance with SEQ ID NO:1 and SEQ ID NO:3.

In a more preferred embodiment the purified nucleic acid segment consists essentially of the nucleotide sequence of SEQ ID NO:1 and SEQ ID NO:3, their complement or the degenerate variants thereof. As used herein, the term "nucleic acid segment" and "DNA segment" are used interchangeably and refer to a DNA molecule which has been isolated free of total genomic DNA of a particular species. Therefore, a "purified" DNA or nucleic acid segment as used herein, refers to a DNA segment which contains a VMP coding sequence yet is isolated away from, or purified free from, total genomic DNA, for example, total cDNA or *borrelia* genomic DNA. Included within the term "DNA segment", are DNA segments and smaller fragments of such segments, and also recombinant vectors, including, for example, plasmids, cosmids, phage, viruses, and the like.

Similarly, a DNA segment comprising an isolated or purified *vls* gene refers to a DNA segment including VMP-related coding sequences isolated substantially away from other naturally occurring genes or protein encoding sequences. In this respect, the term "gene" is used for simplicity to refer to a functional protein, polypeptide or peptide encoding unit. As will be understood by those in the art, this functional term includes both genomic sequences,

cDNA sequences or combinations thereof. "Isolated substantially away from other coding sequences" means that the gene of interest, in this case *vls*, forms the significant part of the coding region of the DNA segment, and that the DNA segment does not contain large portions of naturally-occurring coding DNA, such as large chromosomal fragments or other functional genes or cDNA coding regions. Of course, this refers to the DNA segment as originally isolated, and does not exclude genes or coding regions later added to the segment by the hand of man, nor are other portions or contiguous sequences of naturally occurring DNA excluded.

In particular embodiments, the invention concerns isolated DNA segments and recombinant vectors incorporating DNA sequences which encode a VMP-like protein that includes within its amino acid sequence an amino acid sequence in accordance with SEQ ID NO:2.

Another preferred embodiment of the present invention is a purified nucleic acid segment that encodes a protein in accordance with SEQ ID NO:2, further defined as a recombinant vector. As used herein the term, "recombinant vector", refers to a vector that has been modified to contain a nucleic acid segment that encodes an VMP protein, or a fragment thereof. The recombinant vector may be further defined as an expression vector comprising a promoter operatively linked to said VMP-encoding nucleic acid segment.

A further preferred embodiment of the present invention is a host cell, made recombinant with a recombinant vector comprising an *vls* gene. The recombinant host cell may be a prokaryotic cell. As used herein, the term "engineered" or "recombinant" cell is intended to refer to a cell into which a recombinant gene, such as a gene encoding VMP, has been introduced. Therefore, engineered cells are distinguishable from naturally occurring cells which do not contain a recombinantly introduced gene. Engineered cells are thus cells having a gene or genes introduced through the hand of man. Recombinantly introduced genes will either be in the form of a copy of a genomic gene or a cDNA gene, or will include genes positioned adjacent to a promoter not naturally associated with the particular introduced gene.

In certain embodiments, the invention concerns isolated DNA segments and recombinant vectors which encode a protein or peptide that includes within its amino acid sequence an amino acid sequence essentially as set forth in SEQ ID NO:2. Naturally, where the

DNA segment or vector encodes a full length VMP-like protein, or is intended for use in expressing the VMP-like protein, the most preferred sequences are those which are essentially as set forth in SEQ ID NO:2. It is recognized that SEQ ID NO:2 represents the full length protein encoded by the *vls* gene and that contemplated embodiments include up to the full length sequence and functional variants as well.

The term "a sequence essentially as set forth in SEQ ID NO:2" means that the sequence substantially corresponds to a portion of SEQ ID NO:2 and has relatively few amino acids which are not identical to, or a biologically functional equivalent of, the amino acids of SEQ ID NO:2. The term "biologically functional equivalent" is well understood in the art and is further defined in detail herein, as a gene having a sequence essentially as set forth in SEQ ID NO:1 and that is associated with a *vls* gene in the *Borrelia* family. Accordingly, sequences which have between about 70% and about 80%; or more preferably, between about 85% and about 90%; or even more preferably, between about 90 and 95% and about 99%; of amino acids which are identical or functionally equivalent to the amino acids of SEQ ID NO:2 will be sequences which are "essentially as set forth in SEQ ID NO:2".

In certain other embodiments, the invention concerns isolated DNA segments and recombinant vectors that include within their sequence a nucleic acid sequence essentially as set forth in SEQ ID NO:1 and SEQ ID NO:3. The term "essentially as set forth in SEQ ID NO:1 and SEQ ID NO:3," is used in the same sense as described above and means that the nucleic acid sequence substantially corresponds to a portion of SEQ ID NO:1 and SEQ ID NO:3, and has relatively few codons which are not identical, or functionally equivalent, to the codons of SEQ ID NO:1 and SEQ ID NO:3. The term "functionally equivalent codon" is used herein to refer to codons that encode the same amino acid, such as the six codons for arginine or serine, as set forth in Table 4, and also refers to codons that encode biologically equivalent amino acids.

It will also be understood that amino acid and nucleic acid sequences may include additional residues, such as additional N- or C-terminal amino acids or 5′ or 3′ sequences, and yet still be essentially as set forth in one of the sequences disclosed herein, so long as the sequence meets the criteria set forth above, including the maintenance of biological protein activity where protein expression is concerned. The addition of terminal sequences particularly

applies to nucleic acid sequences which may, for example, include various non-coding sequences flanking either of the 5′ or 3′ portions of the coding region or may include various internal sequences, *i.e.*, introns, which are known to occur within genes.

Excepting intronic or flanking regions, and allowing for the degeneracy of the genetic code, sequences which have between about 70% and about 80%; or more preferably, between about 80%, 85% and about 90%; or even more preferably, between about 90%, 95% and about 99%; of nucleotides which are identical to the nucleotides of SEQ ID NO:1 and SEQ ID NO:3 will be sequences which are "essentially as set forth in SEQ ID NO:1 and SEQ ID NO:3". Sequences which are essentially the same as those set forth in SEQ ID NO:1 and SEQ ID NO:3 may also be functionally defined as sequences which are capable of hybridizing to a nucleic acid segment containing the complement of SEQ ID NO:1 and SEQ ID NO:3 under relatively stringent conditions or conditions of high stringency. Suitable relatively stringent hybridization conditions will be well known to those of skill in the art and are clearly set forth herein, for example conditions for use with Southern and Northern blot analysis, and as described in the examples herein set forth.

Naturally, the present invention also encompasses DNA segments which are complementary, or essentially complementary, to the sequence set forth in SEQ ID NO:1 and SEQ ID NO:3. Nucleic acid sequences which are "complementary" are those which are capable of base-pairing according to the standard Watson-Crick complementary rules. As used herein, the term "complementary sequences" means nucleic acid sequences which are substantially complementary, as may be assessed by the same nucleotide comparison set forth above, or as defined as being capable of hybridizing to the nucleic acid segment of SEQ ID NO:1 and SEQ ID NO:3 under relatively stringent conditions, *i.e.*, conditions of high stringency.

The nucleic acid segments of the present invention, regardless of the length of the coding sequence itself, may be combined with other DNA sequences, such as promoters, polyadenylation signals, additional restriction enzyme sites, multiple cloning sites, other coding segments, and the like, such that their overall length may vary considerably. It is therefore contemplated that a nucleic acid fragment of almost any length may be employed, with the total length preferably being limited by the ease of preparation and use in the intended recombinant DNA protocol. For example, nucleic acid fragments may be prepared which include a short

stretch complementary to SEQ ID NO:1 and SEQ ID NO:3, such as about 10 to 15 or 20, 30, or 40 or so nucleotides, and which are up to 2000 or so base pairs in length. DNA segments with total lengths of about 8000, 7000, 6000, 5000, 4000, 3000, 2000, 1000, 500, 200, 100 and about 50 base pairs in length are also contemplated to be useful.

A preferred embodiment of the present invention is a nucleic acid segment which comprises at least a 14-nucleotide long stretch which corresponds to, or is complementary to, the nucleic acid sequence of SEQ ID NO:1 and SEQ ID NO:3. In a more preferred embodiment the nucleic acid is further defined as comprising at least a 20 nucleotide long stretch, a 30 nucleotide long stretch, 50 nucleotide long stretch, 100 nucleotide long stretch, or at least an 2000 nucleotide long stretch which corresponds to, or is complementary to, the nucleic acid sequence of SEQ ID NO:1 and SEQ ID NO:3. The nucleic acid segment may be further defined as having the nucleic acid sequence of SEQ ID NO:1 and SEQ ID NO:3.

A related embodiment of the present invention is a nucleic acid segment which comprises at least a 14-nucleotide long stretch which corresponds to, or is complementary to, the nucleic acid sequence of SEQ ID NO:1 and SEQ ID NO:3, further defined as comprising a nucleic acid fragment of up to 10,000 basepairs in length. A more preferred embodiment if a nucleic acid fragment comprising from 14 nucleotides of SEQ ID NO:1 and SEQ ID NO:3 up to 5,000 basepairs in length, 3,000 basepairs in length, 2,000 basepairs in length, 1,000 basepairs in length, 500 basepairs in length, or 100 basepairs in length.

Naturally, it will also be understood that this invention is not limited to the particular nucleic acid and amino acid sequences of SEQ ID NO:1 and SEQ ID NO:3. Recombinant vectors and isolated DNA segments may therefore variously include the VMP-like protein coding regions themselves, coding regions bearing selected alterations or modifications in the basic coding region, or they may encode larger polypeptides which nevertheless include VMP-coding regions or may encode biologically functional equivalent proteins or peptides which have variant amino acids sequences.

The DNA segments of the present invention encompass biologically functional equivalent VMP-like proteins and peptides. Such sequences may arise as a consequence of codon redundancy and functional equivalency which are known to occur naturally within

nucleic acid sequences and the proteins thus encoded. Alternatively, functionally equivalent proteins or peptides may be created *via* the application of recombinant DNA technology, in which changes in the protein structure may be engineered, based on considerations of the properties of the amino acids being exchanged. Changes designed by man may be introduced through the application of site-directed mutagenesis techniques, *e.g.*, to introduce improvements to the antigenicity of the VMP-like protein or to test VMP-like mutants in order to examine activity or determine the presence of VMP-like peptide in various cells and tissues at the molecular level.

A preferred embodiment of the present invention is a purified composition comprising a polypeptide having an amino acid sequence in accordance with SEQ ID NO:2. The term "purified" as used herein, is intended to refer to an VMP-related protein composition, wherein the VMP-like protein is purified to any degree relative to its naturally-obtainable state, *i.e.*, in this case, relative to its purity within a eukaryotic cell extract. A preferred cell for the isolation of VMP-like protein is from *borrelia* organisms; however, VMP-like protein may also be isolated from various patient specimens, specimens from infected animals, recombinant cells, tissues. isolated subpopulations of tissues, and the like. as will be known to those of skill in the art, in light of the present disclosure. A purified VMP-like protein composition therefore also refers to a polypeptide having the amino acid sequence of SEQ ID NO:2, free from the environment in which it may naturally occur.

If desired, one may also prepare fusion proteins and peptides, *e.g.*, where the VMP-like protein coding regions are aligned within the same expression unit with other proteins or peptides having desired functions, such as for purification or immunodetection purposes (*e.g.*, proteins which may be purified by affinity chromatography and enzyme label coding regions, respectively).

Turning to the expression of the *vls* gene whether from cDNA based or genomic DNA, one may proceed to prepare an expression system for the recombinant preparation of VMP-like protein. The engineering of DNA segment(s) for expression in a prokaryotic or eukaryotic system may be performed by techniques generally known to those of skill in recombinant expression. For example, one may prepare a VMP-GST (glutathione-S-transferase) fusion

protein that is a convenient means of bacterial expression. However, it is believed that virtually any expression system may be employed in the expression of VMP-like proteins.

VMP-like proteins may be successfully expressed in eukaryotic expression systems. however, the inventors contemplate that bacterial expression systems may be used for the preparation of VMP for all purposes. The cDNA containing *vls* gene may be separately expressed in bacterial systems, with the encoded proteins being expressed as fusions with ß-galactosidase, avidin, ubiquitin, *Schistosoma japonicum* glutathione S-transferase, multiple histidines, epitope-tags and the like. It is believed that bacterial expression will ultimately have advantages over eukaryotic expression in terms of ease of use and quantity of materials obtained thereby.

It is proposed that transformation of host cells with DNA segments encoding VMP-like proteins will provide a convenient means for obtaining a VMP-like protein. It is also proposed that cDNA, genomic sequences, and combinations thereof, modified by the addition of a eukaryotic or viral promoter, are suitable for eukaryotic expression, as the host cell will of course, process the genomic transcripts to yield functional mRNA for translation into proteir.

Another embodiment is a method of preparing a protein composition containing growing recombinant host cell comprising a vector that encodes a protein which includes an amino acid sequence in accordance with SEQ ID NO:2, under conditions permitting nucleic acid expression and protein production followed by recovering the protein so produced. The host cell, conditions permitting nucleic acid expression, protein production and recovery, will be known to those of skill in the art, in light of the present disclosure of the *vls* gene.

2.4 Gene Constructs and DNA Segments

As used herein, the terms "gene" and "DNA segment" are both used to refer to a DNA molecule that has been isolated free of total genomic DNA of a particular species. Therefore, a gene or DNA segment encoding an VMP-like polypeptide refers to a DNA segment that contains sequences encoding an VMP-like protein, but is isolated away from, or purified free from, total genomic DNA of the species from which the DNA is obtained. Included within the term "DNA segment", are DNA segments and smaller fragments of such segments, and also

recombinant vectors, including, for example, plasmids, cosmids, phage, retroviruses, adenoviruses, and the like.

The term "gene" is used for simplicity to refer to a functional protein or peptide encoding unit. As will be understood by those in the art, this functional term includes both genomic sequences and cDNA sequences. "Isolated substantially away from other coding sequences" means that the gene of interest, in this case, a VMP-like protein encoding gene, forms the significant part of the coding region of the DNA segment, and that the DNA segment does not contain large portions of naturally-occurring coding DNA, such as large chromosomal fragments or other functional genes or cDNA coding regions. Of course, this refers to the DNA segment as originally isolated, and does not exclude genes or coding regions, such as sequences encoding leader peptides or targeting sequences, later added to the segment by the hand of man.

2.5 Recombinant Vectors Expressing VMP-like Proteins

A particular aspect of this invention provides novel ways in which to utilize VMP-encoding DNA segments and recombinant vectors comprising *vls* DNA segments. As is well known to those of skill in the art, many such vectors are readily available. one particular detailed example of a suitable vector for expression in mammalian cells is that described in U. S. Patent 5,168,050, incorporated herein by reference. However, there is no requirement that a highly purified vector be used, so long as the coding segment employed encodes a VMP-like protein and does not include any coding or regulatory sequences that would have an adverse effect on cells. Therefore, it will also be understood that useful nucleic acid sequences may include additional residues, such as additional non-coding sequences flanking either of the 5′ or 3′ portions of the coding including, for example, promoter regions, or may include various internal sequences, *i.e.*, introns, which are known to occur within genes.

After identifying an appropriate VMP-encoding gene or DNA molecule. it may be inserted into any one of the many vectors currently known in the art, so that it will direct the expression and production of the VMP-like protein when incorporated into a host cell. In a recombinant expression vector, the coding portion of the DNA segment is positioned under the control of a promoter. The promoter may be in the form of the promoter which is naturally

associated with a VMP-encoding gene, as may be obtained by isolating the 5′ non-coding sequences located upstream of the coding segment or exon, for example, using recombinant cloning and/or PCR™ technology, in connection with the compositions disclosed herein.

In certain embodiments, it is contemplated that particular advantages will be gained by positioning the VMP-encoding DNA segment under the control of a recombinant, or heterologous, promoter. As used herein, a recombinant or heterologous promoter is intended to refer to a promoter that is not normally associated with a *vls* gene in its natural environment. Such promoters may include those normally associated with other *borrelia*-inhibitory polypeptide genes, and/or promoters isolated from any other bacterial, viral, eukaryotic, or mammalian cell. Naturally, it will be important to employ a promoter that effectively directs the expression of the DNA segment in the particular cell containing the vector comprising a *vls* gene or gene segment.

The use of recombinant promoters to achieve protein expression is generally known to those of skill in the art of molecular biology, for example, see Sambrook *et al.*, (1989). The promoters employed may be constitutive, or inducible, and can be used under the appropriate conditions to direct high level or regulated expression of the introduced DNA segment. The currently preferred promoters are those such as CMV, RSV LTR, the SV40 promoter alone, and the SV40 promoter in combination with the SV40 enhancer.

2.6 Methods of DNA Transfection

Technology for introduction of DNA into cells is well-known to those of skill in the art. Five general methods for delivering a gene into cells have been described: (1) chemical methods (Graham and VanDerEb, 1973); (2) physical methods such as microinjection (Capecchi, 1980), electroporation (Wong and Neumann, 1982; Fromm *et al.*, 1985) and the gene gun (Yang *et al.*, 1990); (3) viral vectors (Clapp, 1993; Danos and Heard, 1992; Eglitis and Anderson, 1988); (4) receptor-mediated mechanisms (Wu *et al.*, 1991; Curiel *et al.*, 1991; Wagner *et al.*, 1992); and (5) direct injection of purified DNA into human or animals.

2.7 Liposomes and Nanocapsules

The formation and use of liposomes is generally known to those of skill in the art (see for example, Couvreur *et al.*, 1991 which describes the use of liposomes and nanocapsules in the targeted antibiotic therapy of intracellular bacterial infections and diseases). Recently, liposomes were developed with improved serum stability and circulation half-times (Gabizon and Papahadjopoulos, 1988; Allen and Choun, 1987). The following is a brief description of these DNA delivery modes.

Nanocapsules can generally entrap compounds in a stable and reproducible way (Henry-Michelland *et al.*, 1987). To avoid side effects due to intracellular polymeric overloading, such ultrafine particles (sized around 0.1 mm) should be designed using polymers able to be degraded *in vivo*. Biodegradable polyalkyl-cyanoacrylate nanoparticles that meet these requirements are contemplated for use in the present invention, and such particles may be are easily made, as described (Couvreur *et al.*, 1984; 1988).

Liposomes are formed from phospholipids that are dispersed in an aqueous medium and spontaneously form multilamellar concentric bilayer vesicles (also termed multilamellar vesicles (MLVs). MLVs generally have diameters ranging from 25 mm to 4 mm. Sonication of MLVs results in the formation of small unilamellar vesicles (SUVs) with diameters in the range of 200 to 500 Å, containing an aqueous solution in the core.

In addition to the teachings of Couvreur *et al.* (1991), the following information may be utilized in generating liposomal formulations. Phospholipids can form a variety of structures other than liposomes when dispersed in water, depending on the molar ratio of lipid to water. At low ratios the liposome is the preferred structure. The physical characteristics of liposomes depend on pH, ionic strength and the presence of divalent cations. Liposomes can show low permeability to ionic and polar substances, but at elevated temperatures undergo a phase transition which markedly alters their permeability. The phase transition involves a change from a closely packed, ordered structure, known as the gel state, to a loosely packed, less-ordered structure, known as the fluid state. This occurs at a characteristic phase-transition temperature and results in an increase in permeability to ions, sugars and drugs.

Liposomes interact with cells *via* four different mechanisms: Endocytosis by phagocytic cells of the reticuloendothelial system such as macrophages and neutrophils; adsorption to the cell surface, either by nonspecific weak hydrophobic or electrostatic forces, or by specific interactions with cell-surface components; fusion with the plasma cell membrane by insertion of the lipid bilayer of the liposome into the plasma membrane, with simultaneous release of liposomal contents into the cytoplasm; and by transfer of liposomal lipids to cellular or subcellular membranes, or *vice versa*, without any association of the liposome contents. It often is difficult to determine which mechanism is operative and more than one may operate at the same time.

2.8 Expression of VMP-like Proteins

For the expression of VMP-like proteins, once a suitable (full-length if desired) clone or clones have been obtained, whether they be cDNA based or genomic, one may proceed to prepare an expression system for the recombinant preparation of VMP-like proteins. The engineering of DNA segment(s) for expression in a prokaryotic or eukaryotic system may be performed by techniques generally known to those of skill in recombinant expression. It is believed that virtually any expression system may be employed in the expression of VMP-like proteins.

VMP-like proteins may be successfully expressed in eukaryotic expression systems, however, it is also envisioned that bacterial expression systems may be preferred for the preparation of VMP-like proteins for all purposes. The cDNA for VMP-like proteins may be separately expressed in bacterial systems, with the encoded proteins being expressed as fusions with b-galactosidase, ubiquitin, *Schistosoma japonicum* glutathione S-transferase, green fluorescent protein and the like. It is believed that bacterial expression will ultimately have advantages over eukaryotic expression in terms of ease of use and quantity of materials obtained thereby.

It is proposed that transformation of host cells with DNA segments encoding VMP-like proteins will provide a convenient means for obtaining VMP-like peptides. Both cDNA and

genomic sequences are suitable for eukaryotic expression, as the host cell will, of course, process the genomic transcripts to yield functional mRNA for translation into protein.

It is similarly believed that almost any eukaryotic expression system may be utilized for the expression of VMP-like proteins, *e.g.*, baculovirus-based, glutamine synthase-based or dihydrofolate reductase-based systems could be employed. However, in preferred embodiments, it is contemplated that plasmid vectors incorporating an origin of replication and an efficient eukaryotic promoter, as exemplified by the eukaryotic vectors of the pCMV series, such as pCMV5, will be of most use.

For expression in this manner, one would position the coding sequences adjacent to and under the control of the promoter. It is understood in the art that to bring a coding sequence under the control of such a promoter, one positions the 5' end of the transcription initiation site of the transcriptional reading frame of the protein between about 1 and about 50 nucleotides "downstream" of (*i.e.*, 3' of) the chosen promoter.

Where eukaryotic expression is contemplated, one will also typically desire to incorporate into the transcriptional unit which includes VMP-like protein, an appropriate polyadenylation site (*e.g.*, 5'-AATAAA-3') if one was not contained within the original cloned segment. Typically, the poly A addition site is placed about 30 to 2000 nucleotides "downstream" of the termination site of the protein at a position prior to transcription termination.

Translational enhancers may also be incorporated as part of the vector DNA. Thus the DNA constructs of the present invention should also preferable contain one or more 5' non-translated leader sequences which may serve to enhance expression of the gene products from the resulting mRNA transcripts. Such sequences may be derived from the promoter selected to express the gene or can be specifically modified to increase translation of the RNA. Such regions may also be obtained from viral RNAs, from suitable eukaryotic genes, or from a synthetic gene sequence (Griffiths, *et al.*, 1993).

Such "enhancer" sequences may be desirable to increase or alter the translational efficiency of the resultant mRNA. The present invention is not limited to constructs where the

enhancer is derived from the native 5'-nontranslated promoter sequence, but may also include non-translated leader sequences derived from other non-related promoters such as other enhancer transcriptional activators or genes.

It is contemplated that virtually any of the commonly employed host cells can be used in connection with the expression of VMPs in accordance herewith. Examples include cell lines typically employed for eukaryotic expression such as 239, AtT-20, HepG2, VERO, HeLa, CHO, WI 38, BHK, COS-7, RIN and MDCK cell lines.

It is contemplated that VMP-like protein may be "overexpressed", *i.e.*, expressed in increased levels relative to its natural expression in borrelia cells, or even relative to the expression of other proteins in a recombinant host cell containing VMP-encoding DNA segments. Such overexpression may be assessed by a variety of methods, including radio-labeling and/or protein purification. However, simple and direct methods are preferred, for example, those involving SDS/PAGE and protein staining or Western blotting, followed by quantitative analyses, such as densitometric scanning of the resultant gel or blot. A specific increase in the level of the recombinant protein or peptide in comparison to the level in natural VMP-producing animal cells is indicative of overexpression, as is a relative abundance of the specific protein in relation to the other proteins produced by the host cell and, *e.g.*, visible on a gel.

As used herein, the term "engineered" or "recombinant" cell is intended to refer to a cell into which a recombinant gene, such as a gene encoding a VMP peptide has been introduced. Therefore, engineered cells are distinguishable from naturally occurring cells which do not contain a recombinantly introduced gene. Engineered cells are thus cells having a gene or genes introduced through the hand of man. Recombinantly introduced genes will either be in the form of a cDNA gene (*i.e.*, they will not contain introns), a copy of a genomic gene, or will include genes positioned adjacent to a promoter not naturally associated with the particular introd ed gene.

It will be understood that recombinant VMPs may differ from naturally produced VMP in certain ways. In particular, the degree of post-translational modifications, such as, for

example, lipidation, glycosylation and phosphorylation may be different between the recombinant VMP and the VMP polypeptide purified from a natural source, such as *Borrelia.*

After identifying an appropriate DNA molecule by any or a combination of means as described above, the DNA may then be inserted into any one of the many vectors currently known in the art and transferred to a prokaryotic or eukaryotic host cell where it will direct the expression and production of the so-called "recombinant" version of the protein. The recombinant host cell may be selected from a group consisting of *S. mutans, E. coli, S. cerevisae. Bacillus sp., Lactococci sp., Enterococci sp., or Salmonella sp.* In certain preferred embodiments, the recombinant host cell will have a *rec*A phenotype.

Where the introduction of a recombinant version of one or more of the foregoing genes is required, it will be important to introduce the gene such that it is under the control of a promoter that effectively directs the expression of the gene in the cell type chosen for engineering. In general, one will desire to employ a promoter that allows constitutive (constant) expression of the gene of interest. The use of these constitutive promoters will ensure a high. constant level of expression of the introduced genes. The level of expression from the introduced genes of interest can vary in different clones, probably as a function of the site of insertion of the recombinant gene in the chromosomal DNA. Thus, the level of expression of a particular recombinant gene can be chosen by evaluating different clones derived from each transfection study; once that line is chosen, the constitutive promoter ensures that the desired level of expression is permanently maintained. It may also be possible to use promoters that are subject to regulation, such as those regulated by the presence of lactose analog or by the expression of bacteriophage T7 DNA polymerase.

2.9 Recombinant VMP-like Polypeptides

Recombinant versions of a protein or polypeptide are deemed as part of the present invention. Thus one may, using techniques familiar to those skilled in the art, express a recombinant version of the polypeptide in a recombinant cell to obtain the polypeptide from such cells. The techniques are based on cloning of a DNA molecule encoding the polypeptide from a DNA library, that is, on obtaining a specific DNA molecule distinct from other DNAs.

One may, for example, clone a cDNA molecule, or clone genomic DNA. Techniques such as these would also be appropriate for the production of the VMP-like polypeptides in accordance with the present invention.

2.10 Enhanced Production of VMP-like Proteins

Potential problems with VMP-like proteins isolated from natural sources are low yields and extensive purification processes. An aspect of the present invention is the enhanced production of VMP-like proteins by recombinant methodologies in a bacterial host, employing DNA constructs to transform Gram-positive or Gram-negative bacterial cells. For example, the use of *Escherichia coli* expression systems is well known to those of skill in the art, as is the use of other bacterial species such as *Bacillus subtilis* or *Streptococcus sanguis*.

Further aspects of the invention include high expression vectors incorporating DNA encoding novel *vls*, combinatorial segments and its variants. It is contemplated that vectors providing enhanced expression of VMP in other systems such as *S. mutans* will also be obtainable. Where it is desirable, modifications of the physical properties of VMP may be sought to increase its solubility or expression in liquid culture. The *vls* locus may be placed under control of a high expression promoter or the components of the expression system altered to enhance expression.

In further embodiments, the DNA encoding the VMP-like proteins of the present invention allows for the large scale production and isolation of VMP-like polypeptides. This can be accomplished by directing the expression of the VMP-like polypeptide by cloning the DNA encoding the VMP-like polypeptide into a suitable expression vector. Such an expression vector may then be transformed into a host cell that is able to produce the VMP-like proteins. The VMP-like protein may then be purified, *e.g.*, by means provided for in this disclosure and utilized in a biologically active form. Non-biologically active recombinant VMP-like proteins may also have utility, *e.g.*, as an immunogen to prepare anti-VM antibodies.

2.11 Gene Immunization

Adenovirus is particularly suitable for use as a gene transfer vector because of its mid-sized DNA genome, ease of manipulation, high titer, wide target-cell range, and high infectivity. The roughly 36 kB viral genome is bounded by 100-200 base pair (bp) inverted terminal repeats (ITR), in which are contained *cis*-acting elements necessary for viral DNA replication and packaging. The early (E) and late (L) regions of the genome that contain different transcription units are divided by the onset of viral DNA replication.

The E1 region (E1A and E1B) encodes proteins responsible for the regulation of transcription of the viral genome and a few cellular genes. The expression of the E2 region (E2A and E2B) results in the synthesis of the proteins for viral DNA replication. These proteins are involved in DNA replication, late gene expression, and host cell shut off (Renan, 1990). The products of the late genes (L1, L2, L3, L4 and L5), including the majority of the viral capsid proteins, are expressed only after significant processing of a single primary transcript issued by the major late promoter (MLP). The MLP (located at 16.8 map units) is particularly efficient during the late phase of infection, and all the mRNAs issued from this promoter possess a 5′ tripartite leader (TL) sequence which makes them preferred mRNAs for translation.

In order for adenovirus to be optimized for gene therapy, it is necessary to maximize the carrying capacity so that large segments of DNA can be included. It also is very desirable to reduce the toxicity and immunologic reaction associated with certain adenoviral products.

The large displacement of DNA is possible because the *cis* elements required for viral DNA replication all are localized in the inverted terminal repeats (ITR) (100-200 bp) at either end of the linear viral genome. Plasmids containing ITR's can replicate in the presence of a non-defective adenovirus (Hay *et al.*, 1984). Therefore, inclusion of these elements in an adenoviral vector should permit replication.

In addition, the packaging signal for viral encapsidation is localized between 194-385 bp (0.5-1.1 map units) at the left end of the viral genome (Hearing *et al.*, 1987). This signal mimics the protein recognition site in bacteriophage λ DNA where a specific sequence close to the left end, but outside the cohesive end sequence, mediates the binding to proteins that are

required for insertion of the DNA into the head structure. E1 substitution vectors of Ad have demonstrated that a 450 bp (0-1.25 map units) fragment at the left end of the viral genome could direct packaging in 293 cells (Levrero *et al.*, 1991).

It has been shown that certain regions of the adenoviral genome can be incorporated into the genome of mammalian cells and the genes encoded thereby expressed. These cell lines are capable of supporting the replication of an adenoviral vector that is deficient in the adenoviral function encoded by the cell line. There also have been reports of complementation of replication deficient adenoviral vectors by "helping" vectors, *e.g.*, wild-type virus or conditionally defective mutants.

2.12 VMP-like Variants

VMP-like related proteins and functional variants are also considered part of the invention. Thus it is expected that truncated and mutated versions of VMP-like protein (SEQ ID NO:2) will afford more convenient and effective forms of VMP for treatment regimens. Thus, any functional version of SEQ ID NO:2, such as truncated species or homologs, and mutated versions of VMP-like protein are considered as part of the invention.

Mutagenized recombinant VMPs may have increased potency and prolonged *in vivo* half-life, thereby offering more effective long-term treatments. Novel VMPs thus may be obtained by modifications to the *vls* gene, (such as by site-specific mutagenesis).

Additionally, the 15 silent *vls* cassettes of *B. burgdorferi* may be recombined in numerous combinations, providing for example a cocktail of peptide compositions for use as immunogens and to develop vaccines for use in Lyme disease and related conditions.

2.13 Pharmaceutical Compositions

Pharmaceutical compositions prepared in accordance with the present invention find use in preventing or ameliorating conditions associated with *Borrelia* infections, particularly Lyme disease. Such methods generally involve administering a pharmaceutical composition comprising an effective amount of a VMP-like antigen, such as SEQ ID NO:2 or various

epitopes thereof. Other exemplary compositions may include an effective amount of either a VMP-like variant or a VMP-like encoding nucleic acid composition. Such compositions may also be used to generate an immune response in an animal in such cases where it may be desirable to block the effect of a naturally produced VMP-like protein.

Also included as part of the present invention therefore are novel compositions comprising nucleic acids which encode a VMP-like protein. It will, of course, be understood that one or more than one gene may be used in the methods and compositions of the invention. The nucleic acid delivery methods may thus entail the administration of one, two, three, or more, homologous VMP-encoding genes. The maximum number of genes that may be applied is limited only by practical considerations, such as the effort involved in simultaneously preparing a large number of gene constructs or even the possibility of eliciting an adverse cytotoxic effect.

The particular combination of genes may be two or more distinct genes; or it may be such that a *vls* gene is combined with another gene and/or another protein, cofactor or other biomolecule; a cytokine gene may even be combined with a gene encoding a cell surface receptor capable of interacting with the polypeptide product of the first gene.

In using multiple genes, they may be combined on a single genetic construct under control of one or more promoters, or they may be prepared as separate constructs of the same or different types. Thus, an almost endless combination of different genes and genetic constructs may be employed. Certain gene combinations may be designed to, or their use may otherwise result in, achieving synergistic effects in affording protection against *Borrelia* and/or stimulation of an immune response. Any and all such combinations are intended to fall within the scope of the present invention. Indeed, many synergistic effects have been described in the scientific literature, so that one of ordinary skill in the art would readily be able to identify likely synergistic gene combinations, or even gene-protein combinations.

It will also be understood that, if desired, the nucleic acid segment or gene encoding a VMP-like protein could be administered in combination with further agents, such as, *e.g.*, proteins or polypeptides or various pharmaceutically active agents. So long as the composition comprises a *vls* gene, there is virtually no limit to other components which may also be

included, given that the additional agents do not cause a significant adverse effect upon contact with the target cells or host tissues. The nucleic acids may thus be delivered along with various other agents as required in the particular instance.

2.14 Kits

Therapeutic kits comprising VMP-like peptides or VMP-encoding nucleic acid segments comprise another aspect of the present invention. Such kits will generally contain, in suitable container means, a pharmaceutically acceptable formulation of a VMP-like peptide or a VMP-encoding nucleic acid composition. The kit may have a single container means that contains the VMP composition or it may have distinct container means for the VMP composition and other reagents which may be included within such kits.

The components of the kit may be provided as liquid solution(s), or as dried powder(s). When the components are provided in a liquid solution, the liquid solution is an aqueous solution, with a sterile aqueous solution being particularly preferred. When reagents or components are provided as a dry powder, the powder can be reconstituted by the addition of a suitable solvent. It is envisioned that the solvent may also be provided in another container means.

Kits may also comprise reagents for detecting VMP-like polypeptides, such as required for immunoassay. The immunodetection reagent will typically comprise a label associated with the antibody or antigen, or associated with a secondary binding ligand. Exemplary ligands might include a secondary antibody directed against the first antibody or antigen or a biotin or avidin (or streptavidin) ligand having an associated label. Of course, a number of exemplary labels are known in the art and all such labels may be employed in connection with the present invention. The kits may contain antibody-label conjugates either in fully conjugated form, in the form of intermediates, or as separate moieties to be conjugated by the user of the kit.

The container means will generally include at least one *vial*, test tube, flask, bottle, syringe or other container means, into which the antigen or antibody may be placed, and preferably suitably aliquoted. Where a second binding ligand is provided, the kit will also generally contain a second *vial* or other container into which this ligand or antibody may be

placed. The kits of the present invention will also typically include a means for containing the antibody, antigen, and reagent containers in close confinement for commercial sale. Such containers may include injection or blow-molded plastic containers into which the desired *vials* are retained.

2.15 VMP Antibodies

In another aspect, the present invention contemplates an antibody that is immunoreactive with a polypeptide of the invention. An antibody can be a polyclonal or a monoclonal antibody. In a preferred embodiment, an antibody is a monoclonal antibody. Means for preparing and characterizing antibodies are well known in the art (See, *e.g.,* Howell and Lane, 1988).

Briefly, a polyclonal antibody is prepared by immunizing an animal with an immunogen comprising a polypeptide of the present invention and collecting antisera from that immunized animal. A wide range of animal species can be used for the production of antisera. Typically an animal used for production of anti-antisera is a rabbit, a mouse, a rat, a hamster or a guinea pig. Because of the relatively large blood volume of rabbits, a rabbit is a preferred choice for production of polyclonal antibodies.

Antibodies, both polyclonal and monoclonal, specific for VMP-like polypeptides and particularly those represented by SEQ ID NO:2, variants and epitopes thereof, may be prepared using conventional immunization techniques, as will be generally known to those of skill in the art. A composition containing antigenic epitopes of VMP can be used to immunize one or more experimental animals, such as a rabbit or mouse, which will then proceed to produce specific antibodies against *vls* expression and silent regions. Polyclonal antisera may be obtained, after allowing time for antibody generation, simply by bleeding the animal and preparing serum samples from the whole blood.

To obtain monoclonal antibodies, one would also initially immunize an experimental animal, often preferably a mouse, with a VMP composition. One would then, after a period of time sufficient to allow antibody generation, obtain a population of spleen or lymph cells from the animal. The spleen or lymph cells can then be fused with cell lines, such as human or

mouse myeloma strains, to produce antibody-secreting hybridomas. These hybridomas may be isolated to obtain individual clones which can then be screened for production of antibody to the desired VMP peptide.

Following immunization, spleen cells are removed and fused, using a standard fusion protocol with plasmacytoma cells to produce hybridomas secreting monoclonal antibodies against VMP. Hybridomas which produce monoclonal antibodies to the selected antigens are identified using standard techniques, such as ELISA and Western blot methods. Hybridoma clones can then be cultured in liquid media and the culture supernatants purified to provide the VMP-specific monoclonal antibodies.

It is proposed that the monoclonal antibodies of the present invention will find useful application in standard immunochemical procedures, such as ELISA and Western blot methods, as well as other procedures which may utilize antibody specific to VMP epitopes.

Additionally, it is proposed that monoclonal antibodies specific to the particular polypeptide may be utilized in other useful applications. For example, their use in immunoabsorbent protocols may be useful in purifying native or recombinant VMP species or variants thereof.

In general, both poly- and monoclonal antibodies against VMP may be used in a variety of embodiments. For example, they may be employed in antibody cloning protocols to obtain cDNAs or genes encoding VMP or related proteins. They may also be used in inhibition studies to analyze the effects of VP in cells or animals. Anti-VMP antibodies will also be useful in immunolocalization studies to analyze the distribution of VMP peptides during various cellular events, for example, to determine the cellular or tissue-specific distribution of the VP peptide under different physiological conditions. A particularly useful application of such antibodies is in purifying native or recombinant VMP, for example, using an antibody affinity column. The operation of all such immunological techniques will be known to those of skill in the art in light of the present disclosure.

3.0 BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1A. Correlation of infectivity of *B. burgdorferi* B31 clones 5A1 through 5A10 with presence of a 28-kb linear plasmid (pBB28La). Plasmid profiles of B31 clones as determined by pulse-field gel electrophoresis and ethidium bromide staining. Low- (-) and high- (+) infectivity B31 clones have a virtually identical plasmid banding pattern by this method.

FIG. 1B. Correlation of infectivity of *B. burgdorferi* B31 clones 5A1 through 5A10 with presence of a 28-kb linear plasmid (pBB28La). Hybridization of a DNA blot of the gel shown in FIG. 1A with the pJRZ53 probe. The probe hybridized specifically with a 28-kb plasmid present in all 5 high-infectivity clones but in only 1 of 4 low-infectivity clones. Molecular sizes of the standards are indicated in kilobases, and an asterisk marks the location of pBB28La in the ethidium bromide-stained plasmid profile.

FIG. 2A. Structure of the *vls* locus of *B. burgdorferi* clone B31-5A3. Diagrammatic illustration of the overall arrangement of the *vls* locus in *B. burgdorferi* plasmid pBB28La. Distances from the left telomere are indicated in kb, and the locations of the subtractive hybridization clone pJRZ53 and the λDASII-Bb12 inserts are shown.

FIG. 2B. Structure of the *vls* locus of *B. burgdorferi* clone B31-5A3. Structure of *vlsE*.

FIG. 2C. Structure of the *vls* locus of *B. burgdorferi* clone B31-5A3. Structure of *vlsE*. Nucleotide and predicted amino acid sequences of the allele ***vlsE1*** of the *B. burgdorferi* B31-5A3 *vlsE* gene. The predicted -10 and -35 promoter sequences, the putative ribosome binding site (RBS), and primers used for PCR™ and RT-PCR™ are marked.

FIG. 3A. Sequence similarity of the predicted VlsE sequence (allele VlsE1) with the variable major proteins (Vmps) of *B. hermsii* and the predicted amino acid sequences of the silent *vls* cassettes. Alignment of the predicted amino acid sequence of VlsE (allele VlsE1) with that of Vmp17 (GenBank entry L04788). Identical amino acid residues are indicated by vertical lines (|) and similar residues are marked with colons (:) and periods (.).

FIG. 3B. Sequence similarity of the predicted VlsE sequence (allele VlsE1) with the variable major proteins (Vmps) of *B. hermsii* and the predicted amino acid sequences of the silent *vls* cassettes. Alignment of the deduced peptide sequences of 16 *vls* cassettes. Residues identical to the VlsE cassette region (Vls1) of *B. burgdorferi* are marked as dashes (-); similar amino acids are shown in lower case. Gaps and the predicted stop codons are indicated by dots (.) and asterisk (*), respectively. Variable regions VR-I through VR-VI are shaded.

FIG. 4A. Surface localization of VlsE, as indicated by treatment of intact *B. burgdorferi* with proteinase K. Freshly cultured *B. burgdorferi* B31 clone 5A3 cells were incubated with (+) or without (-) proteinase K at room temperature for 10 min. The proteins of the washed organisms were then separated by SDS-PAGE. The protein blots were reacted with antiserum against the GST-Vls1 fusion protein;

FIG. 4B. Surface localization of VlsE, as indicated by treatment of intact *B. burgdorferi* with proteinase K. Freshly cultured *B. burgdorferi* B31 clone 5A3 cells were incubated with (+) or without (-) proteinase K at room temperature for 10 min. The proteins of the washed organisms were then separated by SDS-PAGE The protein blots were reacted with antiserum against *B. burgdorferi* B31 OspD.

FIG. 4C. Surface localization of VlsE, as indicated by treatment of intact *B. burgdorferi* with proteinase K. Freshly cultured *B. burgdorferi* B31 clone 5A3 cells were incubated with (+) or without (-) proteinase K at room temperature for 10 min. The proteins of the washed organisms were then separated by SDS-PAGE. The protein blots were reacted with monoclonal antibody H9724 against the *B. burgdorferi* flagellin (Fla).

FIG. 5A. Changes in deduced amino acid sequences of VlsE occurring during infection of C3H/HeN mice with *B. burgdorferi* B31-5A3. Flow chart of the overall experimental design.

FIG. 5B. Changes in deduced amino acid sequences of VlsE occurring during infection of C3H/HeN mice with *B. burgdorferi* B31-5A3. Amino acid sequence alignment of the *vlsE* alleles in one clonal population from each of 11 different isolates.

FIG. 5C. Changes in deduced amino acid sequences of VlsE occurring during infection of C3H/HeN mice with *B. burgdorferi* B31-5A3 Amino acid sequence alignment of the *vlsE* alleles in 5 clonal populations from a single ear isolate. In FIG. 5B and FIG. 5C, the deduced amino acid sequences of the mouse isolates were compared with those of the inoculating clone (VlsE1); similarity to this sequence is depicted as described in FIG. 3B. Amino acid residues (EGAIK) encoded by the 17-bp direct repeat are highlighted to indicate the boundaries of the *vls* cassette.

FIG. 6A. Altered VlsE antigenicity of *B. burgdorferi* clones (m1e4A through m8e4A) isolated from C3H/HeN mice 4 weeks post infection. The antigenic reactivties of 9 clones isolated from mice (lanes 109) were compared with those of the parental clone B31-5A3 used for mouse inoculation (lane 11) and the low-infectivity clone B31-5A2 (lane 10), which lacks the plasmid encoding VlsE. Two identical SDS-PAGE western blots were reacted with monoclonal antibody H9724 directed against the *B. burgdorferi* flagellin protein (Fla) as a positive control.

FIG. 6B. Altered VlsE antigenicity of *B. burgdorferi* clones (m1e4A through m8e4A) isolated from C3H/HeN mice 4 weeks post infection and antiserum against the GST-Vls1 fusion protein. Antiserum against the GST-Vls1 fusion protein. Prolonged exposures of the immunoblot against the GST-Vls1 futions protein indicated the presence of weakly reactive bands in all 9 mouse isolates. The relative locations of protein standards are indicated.

FIG. 6C. Reactivity of serum antibodies from a representative *Mus musculus* C3H/HeN mouse with VlsE. An immunoblot of *B. burgdorferi* proteins from the strains indicated and the GST-Vls1 fusion protein were reacted with serum from mouse 1 obtained 28 days after needle inoculation with 10^5 ***B. burgdorferi*** B31, clone 5A3.

FIG. 6D. Reactivity of serum antibodies from a representative *Mus musculus* C3H/HeN mouse with VlsE. An immunoblot of *B. burgdorferi* proteins from the strains indicated and the GST-Vls1 fusion protein were reacted with serum from a Peromyscus leukopcus mouse infected with *B. burgdorferi* B31 via tick-bite. The protein bands corresponding to VlsE and the SGT-Vls1 fusion protein (as determined by reactivity with anti-GST-Vls1 antiserum; data

not shown) are indicated by arrows. The relative locations of protein standards are shown in kilodaltons.

FIG. 6E. Reactivity of serum antibodies from a representative *Mus musculus* C3H/HeN mouse with VlsE. An immunoblot of *B. burgdorferi* [proteins from the strains indicated and the GST-Vls1 fusion protein were reacted with serum from an early stage Lyme disease patient. The protein bands corresponding to VlsE and the GST-Vls1 fusion protein (as determined by reactivity with anti-GST-Vls1 antiserum; data not shown) are indicated by arrows. The relative locations of protein standards are shown in kilodaltons.

FIG. 7. Proposed model for genetic and antigenic variation at the *vls* locus. Recombination of segments of the silent *vls* cassettes *vls7* and *vls4* into the *vls1* cassette of *B. burgdorferi* B31-5A3 *vlsE* gene is shown. A series of similar recombination events would generate unique *vlsE* alleles consisting of a mosaic of segments from several different silent *vls* cassettes.

4.0 DETAILED DESCRIPTION OF ILLUSTRATIVE EMBODIMENTS

The present work discloses the identification and characterization of an elaborate genetic system in the Lyme disease spirochete *Borrelia burgdorferi* that promotes extensive antigenic variation of a surface-exposed lipoprotein, VlsE. A 28-kilobase linear plasmid of *B. burgdorferi* B31 (pBB28La) was found to contain a *vmp*-like sequence (*vls*) locus that closely resembles the variable major protein (*vmp*) system for antigenic variation of relapsing fever organisms. Portions of several of the 15 non-expressed (silent) *vls* cassette sequences located upstream of *vlsE* recombined into the central *vlsE* cassette region during infection of C3H/HeN mice, resulting in antigenic variation of the expressed lipoprotein. The resulting combinatorial variation will potentially produce millions of unique antigenic variants and thereby contribute to immune evasion, long-term survival, and pathogenesis in the mammalian host.

An infectivity-associated 28-kb linear plasmid, pBB28La, in *B. burgdorferi* B31 by subtractive hybridization has been identified. DNA sequence analysis of cloned fragments from this plasmid revealed the *vls* locus consisting of 15 silent *vls* cassettes and an expressed *vlsE*

gene. Subsequent infection studies demonstrated that promiscuous recombination occurs at the *vlsE* site in C3H/HeN mice. Although the *vls* locus has been characterized thoroughly only in one clonal population of *B. burgdorferi* B31, Southern hybridization results indicate that this locus is present in infectious strains of three well-defined Lyme disease *Borreliae* genospecies (*B. burgdorferi*, *B. afzelii*, and *B. garinii*), despite the overall genetic heterogeneity among these organisms (Casjens *et al.*, 1995; Xu and Johnson, 1995).

The *vls* locus resembles the *vmp* system of *B. hermsii* in both sequence and genetic organization. There is some sequence homology between these two systems, particularly between the *vlsE* and large *vmp* genes. This is exemplified by direct sequence comparison between *vlsE* and *vmp17*, which have homology throughout their predicted amino acid sequences (FIG. 3A). The *vlsE* and silent *vls* cassettes also have a closer degree of homology to small *vmps* and *B. burgdorferi ospC* genes. Additionally, both the *vls* and *vmp* systems have a single expression site encoding a surface-localized lipoprotein, as well as multiple unexpressed sequences (Plasterk *et al.*, 1985; Barbour *et al.*, 1991a). Finally, the expression sites for both systems are located near one of the telomeres of their respective linear plasmids (Kitten and Barbour, 1990; Barbour *et al.*, 1991b). These observations suggest that the *vls* locus may provide the Lyme disease *Borreliae* with the capability of antigenic variation analogous to the *vmp* system of *B. hermsii* (Barbour, 1993). The above similarities also indicate that the *vlsE* gene, silent *vls* cassettes, and large *vmp* genes of relapsing fever organisms, all evolved from a common ancestral gene. Their relatively high G+C compositions (*e.g.* 45% for *vlsE* and 37% for *vmp17*) when compared with *Borrelia* G+C content (~28%) are also consistent with this evolutionary relationship, and further suggest the possibility of lateral transfer from other organisms.

There are several differences between the *vls* and *vmp* systems. First, *B. hermsii* possesses at least two *vmp*-containing linear plasmids (Meier *et al.*, 1985; Plasterk *et al.*, 1985), whereas only one *vls*-containing linear plasmid was detected in Lyme disease *Borreliae* under hybridization conditions (FIGs. 1A and1B). Second, the silent *vmp* genes are separated by intergenic noncoding regions and arranged in either orientation (Barbour *et al.*, 1991a), but the silent *vls* cassettes are organized head-to-tail as a single open reading frame throughout almost the entire region (FIGs. 2A and 2B). Third, the silent *vmp* genes lack promoter sequences, but

most encode complete or nearly complete open reading frames with their own ribosome-binding sites (Barbour *et al.*, 1991a). On the other hand, the *vls* cassettes represent only the central third of the expression site. Lastly, each phase of *B. hermsii* infection is caused predominantly by organisms expressing a single *vmp* allele (Meier *et al.*, 1985; Plasterk *et al.*, 1985), whereas a high degree of *vlsE* allelic variation occurs among organisms isolated even from a small ear biopsy specimen during *B. burgdorferi* infection (FIGs. 5A, 5B and 5C).

The sequence changes at the *vlsE* site may result from genetic recombination with sequences from the silent *vls* cassettes. Despite considerable sequence variations within the *vls* region of different *vlsE* alleles, the sequence examined outside the 17-bp direct repeats remained unchanged (FIG. 5B and FIG. 5C). Within the *vls* region, the changes are not random but are clustered predominantly in six highly variable regions found in 15 silent *vls* cassettes (FIG. 3B). Nearly all of the sequence variations observed in the mouse isolates are identical to portions of the silent *vls* cassettes, although the combinations of the sequence variations made each of these alleles unique.

The inventors have shown that *B. burgdorferi* undergoes an unusual type of genetic variation (FIG. 7): (i) the *vls* cassettes contain conserved and variable regions; (ii) the conserved sequences facilitate recombination between the expressed and silent *vls* sequences, probably by a non-reciprocal gene conversion mechanism; (iii) the conserved 17-bp direct repeat sequences may be involved in alignment of the *vls* sequences during recombination or in binding of proposed site-specific recombinase(s); (iv) through multiple recombination events, portions of the expression site are replaced by segments from several silent *vls* cassettes, resulting in a vast array of potential *vlsE* alleles; and (v) the site-specific mechanism is activated *in vivo*, resulting in a high rate of recombination. Since both the *vlsE* and silent *vls* cassettes are located on the same linear plasmid, pBB28La, in *B. burgdorferi* (FIG. 2A), intraplasmic recombination is likely to be involved. However, it is also possible that interplasmic recombination of multiple copies of the pBB28La plasmid are present in each organism, as shown with the *vmp*-encoding plasmids of *B. hermsii* (Kitten and Barbour, 1992).

Genetic variation involved in multi-gene families has been described in several other pathogenic microorganisms (Borst and Geaves, 1987; Borst *et al.*, 1995; Donelson, 1995). In the context of combinatorial recombination, the genetic variation at the *vlsE* site is similar to

that of the pilin-encoding genes of N. Gonorrhoeae (Seifert and So, 1988). The gonococcal pilus is primarily composed of repeating subunits of an 18-kilodalton pilin protein and is required for adherence of the bacterium to a variety of human cells (Swanson and Koomey, 1989). While the complete pilin genes are expressed only at two expression sites (*pilE1* and *pilE2*), multiple silent copies (*pilS*) containing portions of the pilin genes are found over a wide range on the gonococcal chromosome (Haas and Meyer, 1986). Through multiple combinatorial recombination events, a single gonococcal clone expressing one pilin stereotype can give rise to a large number of progeny that express antigenically distinctive pilin variants (Meyer *et al.*, 1982; Hagblom *et al.*, 1985; Segal *et al.*, 1986). The recombination between the expression and silent loci occurs predominantly through a non-reciprocal gene conversion mechanism (Haas and Meyer, 1986; Koomey *et al.*, 1987).

The coding sequences of the Neisseria pilin variants are divided into constant, semi-variable, and hypervariable regions (Haas and Meyer, 1986), which are analogous to the conserved and variable regions of the *vls* cassettes (FIG. 3B, FIG. 5B and FIG. 5C). The constant regions and a conserved DNA sequence (Sma/Cla repeat) located at the 3' end of all pilin loci are though to pair the regions involved in recombination events (Wainwright *et al.*, 1994). In this context, the 17-bp direct repeats (FIG. 2C) and the conserved regions (FIG. 3B) of the *vls* cassettes may play a similar role in recombination events. The silent loci of gonococcal pilin genes contain different regions of the complete pilin genes (Haas and Meyer, 1986), whereas the silent *vls* cassettes of *B. burgdorferi* represent only the central cassette region of the *vlsE* gene (FIG. 3B).

Non-reciprocal recombinations also occurs between the expressed and the silent genes encoding variant surface glycoproteins (Vsgs) in African trypanosomes (Donelson, 1995). Based on similarities between the *vls* locus and the multi-gene families of the other pathogenic microorganisms, it is likely that a unidirectional gene conversion mechanism is also active in the *vls* locus. However, there is not as yet any data regarding the preservation of the silent *vls* cassettes, and the exact mechanism of *vls* recombination remains to be determined.

There is strong evidence that genetic variation at the *vls* locus generates antigenic variation. The prolific recombination at the *vlsE* site in C3H/HeN mice supports the possibility of antigenic variation in Lyme diseases caused by *Borreliae*. The decreased reactivity to

antibody against the parental Vls1 cassette region among the clonal populations of mouse isolates demonstrates that genetic variation at the *vlsE* site resulted in changes in antigenicity of the VlsE variants (FIG. 6B). Finally, C3H/HeN mice infected with *B. burgdorferi* produced strong antibody responses against the parental VlsE protein, but consistent with the results obtained with the antibody against the GST-Vls1 fusion protein, the same antisera had decreased reactivities with some of the VlsE variants isolated from mice (FIG. 6C). Since VlsE is a surface-exposed lipoprotein, as indicated by proteinase K digestion (FIGs. 4A and 4B) and [^{3}H]-palmitate radiolabeling studies, this proposed antigenic variation may allow Lyme disease *Borreliae* to survive immune attack targeted against VlsE.

Variation of *B. burgdorferi* surface proteins such as VlsE may also affect the organism's virulence and its ability to adapt to different micro-environments during infection of the mammalian host. Recent studies of a *Borrelia turicatae* mouse infection model that resembles Lyme disease showed that one serotype expressing VmpB exhibited more severe arthritic manifestations, whereas another expressing VmpA had more severe central nervous system involvement (Cadavid *et* al, 1994). The numbers of *Borreliae* present in the joints and blood of serotype B-infected mice were much higher than those of mice infected with serotype A, consistent with a relationship between Vmp serotype and disease severity (Pennington *et* al, 1997). Antigenic variation of Neisseria pilin (Lambden *et al.*, 1980; Rudel *et al.*, 1992; Nassif *et al.*, 1993; Jonsson *et al*, 1994) and Opa proteins (Kupsch *et al*, 1993) is known to affect adherence of the organisms to human leukocytes and epithelial cells.

The importance of the *vls*-containing plasmid, pBB28La, during infection is supported by the following evidence: (i) all high-infectivity clones and strains tested thus far contain the *vls*-containing plasmid pBB28La, and loss of this plasmid correlates with a decrease in infectivity (FIG. 1B); (ii) pBB28La was maintained in all animal isolates tested thus far; and (iii) the *vls* sequences are preserved among three Lyme disease genospecies despite their genetic heterogeneity (Casjens *et al.*, 1995) and diversity in plasmid profiles (Xu and Johnson, 1995). On the other hand, *B. burgdorferi* clones with or without plasmid pBB28La showed similar growth rates in culture medium. In addition, pBB28La is readily lost during *in vitro* subcultures as early as passage 5. Therefore, presence of pBB28La appears to have little if any effect on *in vitro* growth, yet has a profound effect on the ability to infect mammalian host.

VlsE (or, potentially, other genes encoded by pBB28La) appears to have another important but undefined function which is unrelated to antigenic variation. Low-infectivity clones lacking the *vls*-encoding plasmid pBB28La do not propagate in severe combined immunodeficiency (SCID) mice, indicating that the required factor(s) provides an important function unrelated to evasion of the adaptive immune system. Also, *in vivo* selection against *Bb* clones lacking pBB28La appears to occur early in infection (within the first week), before the adaptive immune response would be expected to exert significant selection pressure. Therefore, it is likely that VlsE plays an important role in some aspect of infection (*e.g.* colonization, dissemination, adherence, extravasation, evasion of innate immune mechanisms, or nutrient acquisition), and that antigenic variation merely permits surface expression of this protein without leading to elimination of the bacteria by the host's immune response. Retention of this activity would require that the variation in amino acid sequences would not interfere with the active site(s) of the protein; this requirement may explain the existence of highly conserved regions at the N- and C-termini and within the *vls* cassette. Sequence variation as a mechanism of maintaining surface protein function in the face of a hostile immune response may be a strategy common to pathogenic microorganisms.

4.1 Antigenic variation in B. hermsii

A complex antigenic variation mechanism has been characterized in *Borrelia hermsii*, a relative of *B. burgdorferi* that causes relapsing fever (Balmelii and Piffatetti, 1996; Barbour, 1993; Donelson, 1995). Surface-exposed lipoproteins called variable major proteins (Vmps) are encoded by homologous genes located in 28- to 32-kb linear plasmids with covalently closed telomeres (Barbour and Garon, 1987; Kitten and Barbour, 1990). The *vmp* genes have been subdivided into two groups: small and large (Restrepo *et al.*, 1992). Large *vmp* genes such as *vmp*7 and *vmp*17 and small *vmp* genes such as *vmp*1 and *vmp*3 are approximately 1 kb and 0.6 kb in size, respectively. Each organism contains both small and large *vmp* genes in a unexpressed (silent) form in the so-called storage plasmids (Plasterk *et al.*, 1985). Only one *vmp* gene located near one of the telomeres of a different plasmid (called the expression plasmid) is expressed in each organism (Kitten and Barbour, 1990; Barbour *et al.*, 1991a). Antigenic variation occurs when the expressed *vmp* is replaced completely or partially by one of the silent *vmp* genes at the telomeric expression site through interplasmic recombination (Meier

et al., 1985; Plasterk *et al.*, 1985; Barbour *et al.*, 1991b), intraplasmic recombination (Restrepo *et al.*, 1994), and post-switch rearrangement (Restrepo and Barbour, 1994). The antigenic switch occurs spontaneously at a frequency of 10^{-3} to 10^{-4} per generation (Stoenner *et al.*, 1982).

4.2 Identification of *vls*

A genetic locus (called *vmp*-like sequence or *vls*) has been identified and characterized in *B. burgdorferi* that surprisingly resembles the *vmp* system of *B. hermsii*. A *vls* expression site (*vlsE*) and 15 additional silent *vls* cassettes were identified on a 28-kb linear plasmid (designated pBB28La). The presence of pBB28La correlates with the high-infectivity phenotype in *B. burgdorferi* sensu lato strains tested. *vlsE*, located near a telomere of pBB28La, encodes a surface-exposed lipoprotein. Examination of ear and blood isolates from C3H/HeN mice infected 4 weeks previously with B31 clone 5A3 demonstrated the occurrence of promiscuous recombination at the *vlsE* site, such that each of *B. burgdorferi* clones examined was unique and appeared to have undergone multiple recombination events with portions of the silent *vls* cassettes. The resultant VlsE variants exhibited a decreased reactivity to antiserum directed against the parental Vls1 cassette region. This elaborate genetic system permits combinatorial antigenic variation of *vlsE* in the mammalian host, thereby contributing to evasion of the immune response and long-term survival in the mammalian host.

The present invention illustrates the rapid occurrence of promiscuous recombination at the *vls* expression site (*vls*E), resulting in a combinatorial form of genetic and antigenic variation at the *vls*E site. Antigenic variation at the *vls* site has been detected using an *in vivo* selection approach.

The sequence variation appears to lead to significant antigenic variation. Rabbit antiserum raised against a *vls1*-GST fusion protein reacted strongly with the original *B. burgdorferi* clone (B31 5A3), but did not react with several of the clones reisolated from mice 4 weeks post infection.

B. burgdorferi induces a site-specific recombination mechanism during infection of the mammalian host. The *vlsE* cassette sequence in each of the mouse isolates is unique. At the nucleotide level each *vlsE* cassette is comprised of regions identical to several of the silent *vls*

cassettes. This promiscuous recombination of silent *vls* cassette segments causes a combinatorial diversity at the *vlsE* expression site, similar to the diversity possible in the immunoglobulin and T cell-receptor variable regions. In contrast, antigenic variation in relapsing fever organisms usually involves replacement of the entire gene at the expression site with one of the 'silent' VMP genes. Moreover, a single VMP serotype is predominant during each relapse.

This mechanism of genetic switching appears to be different from any other antigenic variation mechanism described in bacteria or protozoa and has important implications in Lyme disease. By combining different regions of the silent *vls* cassettes, it is possible for many different *VlsE* ' serotypes' to coexist in the same patient. It may be impossible for the host to mount a protective response against any one of these clonal populations, because of the small number of each type. Even mounting a response against one serotype would not protect against rapidly evolving, new serotypes. The fact that *B. burgdorferi* has evolved such an elaborate mechanism for varying the sequence of *VlsE* indicates the importance of the protein in pathogenesis and/or immune evasion.

The present invention discloses a repetitive DNA sequence ~500 bp in length which is present in multiple, nonidentical copies in a 28 kb linear plasmid of infectious *Borrelia burgdorferi*, the causative agent of Lyme disease. These DNA sequences encode polypeptides which have sequence similarity to the Variable Major Proteins (VMPs) of relapsing fever *Borreliae* (such as *B. hermsii*). VMPs are highly antigenic surface proteins which the relapsing fever *Borreliae* are able to change through a genetic recombination mechanism, thereby evading the immune response. Antibodies against a particular VMP are protective, resulting in rapid clearance of bacteria of the corresponding serotype. In *B. burgdorferi*, VMP-like sequences (*vls*) are present on a 28 kb linear plasmid, and this plasmid appears to encode virulence factor(s) required for infectivity. The sequence of a 16 kb region of this plasmid contains at least 20 copies of the VMP-like sequence.

The inventors have identified genes and gene products that appear to be important in the infectivity and pathogenesis of *B. burgdorferi*. In previous studies (Norris *et al.*, 1995), it was shown that clonal populations of *B. burgdorferi* isolated after 5 to 15 *in vitro* passages varied significantly in their infectivity in the C3H/HeN mouse model. So-called high-infectivity and

low-infectivity clones differed by 500-fold in their median infectious dose (1.8×10^2 vs. 1×10^5), yet exhibited no obvious differences in terms of protein content (as determined by two dimensional gel electrophoresis and silver staining) or plasmid content (determined by agarose gel electrophoresis and ethidium bromide staining). However, by using subtractive hybridization between DNA of high- and low-infectivity organisms, specific sequences that differed between the two types have been identified. These sequences have been characterized as VMP-like sequences (*vls*), now identified for the first time in *B. burgdorferi*.

In initial studies, high-passage (HP) and low-passage (LP) uncloned populations of *B. burgdorferi* strain B31 were used as a source of DNA for subtractive hybridization. HP B31 was cultured *in vitro* for ~1,000 passages and found to be noninfectious, whereas LP B31 passages *in vitro* ~5 times remains infectious in the C3H/HeN mouse model. The plasmid DNA of each strain was purified. The DNA of HP B31 was randomly sheared by ultrasonication, whereas the DNA of LP B31 was digested to completion with the restriction enzyme *Sau*3AI. The DNA of the two strains was denatured by heating to 100°C, mixed at a ratio of 50:1 HP DNA to LP DNA, and allowed to hybridize with the sheared HP DNA; as a result, the *Sau*3AI restriction sites were not regenerated. Unique segments of the LP DNA tend to hybridize with the complementary LP DNA strand, and the *Sau*3AI "sticky ends" are regenerated. A portion of the hybridized mixture was ligated into pBluescript II SK- (Stratagene) that had been treated previously with *Bam*HI and alkaline phosphatase. The ligated preparation was used to transform *E. coli* XL-1 Blue cells, and transformants were selected by plating the bacteria on Luria broth (LB) agar plates containing ampicillin and isopropyl thiogalactopyranoside (IPTG) and 5-bromo-4-chloro-3-indolyl-D-galactoside (X-gal). Any resulting white colonies (*E. coli* containing a plasmid with a DNA insert) were selected for further study and sequence analysis.

One of the resulting recombinant plasmids, designated pJRZ53, contained DNA encoding a single, contiguous open reading frame (ORF) 562 bp in length. The deduced amino acid sequence of this ORF had significant homology with Vmp proteins of *B. hermsii*, most notably Vmp17, Vmp21, Vmp7, and Vmp25 (27.2 to 20.3% identity, 50.0 to 56.8% similarity). Hybridization of pJRZ53 with Southern blots of *B. burgdorferi* plasmids showed that this VMP-like DNA sequence was localized on a 28 kb linear plasmid.

Additional DNA recombinants containing sequences hybridizing with pJRZ53 were derived from a *Pst*I library of *B. burgdorferi* B31 plasmid DNA. *B. burgdorferi* B31 plasmid DNA was treated with several restriction enzymes to determine the best combination for cloning a larger fragment containing the pJYZ53 sequence. Surprisingly, numerous bands hybridizing with the pJRZ53 probe were present in DNA digested with *Pst*I, *Rsa*I, *Sau*3AI and other enzymes. This result demonstrated that multiple sequences resembling the pJRZ53 insert were present in the 28 kb plasmid.

Several additional recombinant clones were obtained by treating *B. burgdorferi* plasmid DNA with *Pst*I, ligating the fragments into pBluescript II SK-, transforming *E. coli* with the resulting recombinants, and screening the library for hybridization with the pHRZ53. Sequence determinations of these recombinant clones confirmed the presence of multiple sequences that were highly homologous but nonidentical to the pHRZ53 sequence. This homology at the DNA and protein levels is exemplified by the comparison of the two contiguous repeats found in pHRZ53-31, an independently derived 1143 bp *Pst*I clone that overlaps the pJRZ53 sequence. Alignment of the 5' and 3' regions of pJRZ53-31 shows highly homologous repeats in the DNA sequence of recombinant clone pJRZ53-31. The DNA sequence from the 5' (nt 1-578) and 3' (nt 579-1143) regions were aligned using the GCG program GAP. There is 93% identity between the 5' and 3' regions. The deduced amino acid sequences of the DNA regions were aligned using the GCG program GAP. The overall sequence similarity and identity are 92% and 85%, respectively.

Subsequent studies used clonal populations of *B. burgdorferi* obtained by subsurface colony formation of passage 5 organisms on agar plates (Norris *et al.*, 1995). These clones were characterized in terms of infectivity, and were subdivided into high-infectivity and low-infectivity phenotypes. pJRZ53 hybridized with a 28 kb band in 6/9 B31 clones and 7/10 Sh2-2-82 clones. All 12 highly infectious clones contained the plasmid, whereas 6 of 7 low infectivity clones lacked the plasmid (Table 2). There was correlation of the presence of a 28 kb plasmid containing the pJRZ53 sequence with infectivity of *B. burgdorferi* clonal populations. Genomic DNA preparations from 10 Sh2-2-82 clones and 9 B31 clones (Norris *et al.*, 1995) were subjected to pulsed field gel electrophoresis and transferred to nylon membranes. The 562 bp pJRZ53 insert labeled with ^{32}P was hybridized to the Southern blots. Controls consisted of

uncloned populations of high-passage, noninfectious B31 (-) and low-passage, infectious B31 (+). All high infectivity clones (+) possessed a 28 kb plasmid that hybridized with pJRZ53, whereas only 1/7 low-infectivity clones (-) had the plasmid.

Thus there is a strong correlation between the presence of the 28 kb plasmid and infectivity. Plasmid profiles in the same gels used for Southern blot hybridization did not reveal any difference in ethidium bromide staining in the region of the 28 kb plasmid, due to the presence of several other comigrating plasmids. The one low infectivity clone that contained the plasmid may lack a functional gene or genes encoding other virulence factors.

TABLE 2

Correlation Between Infectivity and the Presence of the 28 kb Linear Plasmid Hybridizing with the pJRZ53 Sequence

Clonal Populations	Number of clones containing the 28 kb plasmid/total
B31, high infectivity	5/5
B31, low infectivity	1/4
Sh2, high infectivity	7/7
Sh2, low infectivity	0/3

A map of the 28 kb linear plasmid (designated pBb28L) showed a 16 kb fragment of pBb28L of the bacterial clone B31-5A3 had been cloned into the vector lambda Dash II (Stratagene, LaJolla, CA). Briefly, a preparation of plasmid DNA was treated with S1 nuclease to disrupt the covalently closed ends (telomeres) of the linear plasmids. After treatment with Klenow fragment of DNA polymerase, an oligonucleotide linker containing an *Eco*RI site was ligated onto the ends. The preparation was then treated with *Eco*RI (to cleave the DNA both at the linker and at a previously mapped internal *Eco*RI site) and ligated into the *Eco*RI site of lambda Dash II. Clones containing the pJRZ53 sequence were identified by hybridization, and included two overlapping clones 12 kb and 16 kb in length. Partial sequence analysis of the

sequence of the 16 kb fragment revealed the presence of at least 17 VMP-like sequences within this region.

Over 9,500 bp of DNA sequence from the *B. burgdorferi* DNA insert in Lambda DashII Clone 12-1 have been obtained. This sequence was determined through automated sequence analysis (using T3, T7, and internal primers) of DashII 12-1 itself, as well as of portions of 12-1 cloned into pBluescript using fragments obtained by random DNaseI-mediated cleavage or by digestion with *Pst*I or *Rsa*I. Sequences were assembled using the GCG program Gelassemble and have an average redundancy of ~4 fold. The high degree of sequence identity among different regions required careful verification of sequence differences and manual alignment of sequences in some instances.

The sequences obtained indicate the presence of at least one open reading frame representing an expression site (*vlsE1)* and at least 16 additional nonidentical, apparently 'silent' (nonexpressed) *vls* cassettes. A consensus ribosome binding site (RBS, underlined) is located 8 nucleotides (nt) upstream of the predicted translational start site at nt 75-77. The predicted product, VlsE1, has a molecular weight of 35,881 kDa. The first 19 amino acids of the predicted N-terminus contain a possible signal peptide sequence, with a motif of a charged N-terminus, a hydrophobic region, and a potential signal peptidase II cleavage site (FINC, double-underlined) resembling those found in other *Borrelial* lipoproteins. The predicted polypeptide size after cleavage at this site is 33,957 kDa, and the predicted isoelectric point is 7.3. Except for the signal peptide, the predicted protein is largely hydrophilic. The putative stop codon is located at nt 1143-1145, only 82 nt from the telomeric end of pBb28L.

Expression of *vlsE1* in the high infectivity *B. burgdorferi* B31 Clone 3 was verified by Northern blot analysis and reverse transcriptase polymerase chain reaction (RT-PCR™). Hybridization of radiolabelled pJRZ53 insert to blots of RNA separated by agarose electrophoresis indicated the presence of a transcript containing a homologous sequence. For RT-PCR™, primers corresponding to nt 835-857 (plus strand) and nt 1010-1032 (minus strand) of the sequence in FIG. 2C were constructed. The minus strand primer was used in combination with AMV reverse transcriptase and RNA isolated from B31 clone 3 to produce a cDNA product. The cDNA was then amplified by standard PCR™ using the plus and minus primers, resulting in a 198 bp product detectable by ethidium bromide staining of agarose gels.

The PCR™ product was ligated into the pCRII vector (Invitrogen), and three independently-derived clones yielded sequences identical to that shown in SEQ ID NO:1 and SEQ ID NO:3. Control preparations consisted of reactions identical except for the omission of reverse transcriptase; no product was detected. This result demonstrated that *vlsE1* is transcribed in *B. burgdorferi* B31 Clone 3 organisms.

The DNA sequence of a proposed 'storage site' contains at least 15 contiguous copies of the *vls* sequence of SEQ ID NO:1 and SEQ ID NO:3. The beginning and end of each *vls* 'cassette' was selected to match the repetitive sequence (***vls1***) in *vlsE1*. The *vls* cassettes identified thus far range from 474 to 582 nt in length; length variation is primarily due to short insertions or deletions in multiples of three nucleotides, indicating selective preservation of the open reading frames. Longer deletions are seen in *vls7*, *vls8*, and *vls10*. *vls14* and *vls16* each contain one frameshift, and *vls11* contains one stop codon. Otherwise, the 7766 bp sequence of SEQ ID NO:1 and SEQ ID NO:3 represents one contiguous open reading frame.

The *vls* cassettes exhibit a remarkable degree of sequence conservation at both the DNA and encoded amino acid levels. see FIGs . Nucleotide sequences of *B. burgdorferi* B31 *vls* were aligned using the GCG program PILEUP. *vls1* corresponds to nt 420-1003 in the sequence of FIG. 4 and FIG. 5. When compared to *vls1* using GCG program GAP, the *vls* sequences have 90.0% to 96.1% nucleotide sequence identity (FIG. 6), and 76.9% to 91.4% predicted amino acid sequence identity (FIG. 7). None of the *vls* copies identified thus far have complete sequence identity, but all are closely related.

Table 3 shows the vls segments identified and indicates the positions at which the segments may be found as part of SEQ ID NO:1 and SEQ ID NO:3. Repeat recombinant segments are identified as "repeats".

TABLE 3

CASSETTE (vls)	POSITION IN SEQ ID NO: 3	REPEAT position in SEQ ID NO:3
vls 2	<205>-711 (truncated at 5' end)	711-727
vls 3	712-1293	1293-1309
vls 4	1294-1869	1869-1885
vls 5	1870-2439	2439-2456
vls 6	2440-3009	3009-3025
vls 7	3010-3483	3483-3499
vls 8	3484-3990	3990-4006
vls 9	3991-4548	4548-4557
vls 10	4549-5058	5058-5074
vls 11	5059-5652	5652-5668
vls 12	4653-6219	6219-6253
vls 13	6220-6789	6789-6805
vls 14	6846-7373	7373-7389
vls 15	7274-7946	7946-7962
vls 16	7947-8000	

The degree of sequence similarity between the VMP-like sequences and *B. hermsii* VMP proteins were exemplified by an alignment of the predicted translation product of *vlsE1* with some of the most similar VMP sequences (*vmp 17, vmp 21, vmp7*). Regions of similarity are interspersed among areas of low sequence identity. The G+C contents of the *B. burgdorferi* VMP-like sequences are quite high (*e.g.*, 49.9% for pJRZ53-31) as compared to the genomic *B. burgdorferi* G+C content (27 to 30%) or that of *B. hermsii* VMP genes (*e.g.*, 37% for *vmp17*). The sequence similarity at the protein level may be due to divergent or convergent evolution. It is also possible that the VMP-like sequences were acquired from another organism, given the different G+C content.

Alignment of either the DNA sequences or the deduced amino acid sequences of the open reading frames reveals the presence of both conserved and variable regions of the repetitive sequence. The conserved sequences may represent 'framework' regions important in the overall structure of the polypeptides, whereas the variable sequences may produce different

epitopes. It is contemplated that protective antibodies can be produced against either the conserved or variable portions of the putative amino acid sequences.

The expressed copy of vls (vlsE1) has been identified and sequenced. A segment of the *vlsE* gene corresponding to the cassette region has been subcloned into the pGEX-2T expression vector, and the resulting GST-vls1 fusion protein product produced and purified. Antibodies against the recombinant protein have been used for identification of the native protein in SDS-PAGE and two dimensional gel electrophoresis patterns of *B. burgdorferi* proteins by immunoblotting. Infected patients and animals produced antibodies against the protein which were detected by immunoblot analysis using the recombinant protein as antigen (FIGs. 6C, 6D and 6E). In addition, the purified recombinant protein may be used for immunization of mice and other animals to determine whether antibodies or cellular responses against the protein are protective against infection with *B. burgdorferi* and other Lyme disease *Borreliae*. Such animal studies would determine the feasibility of vaccination of humans and animals with Vls protein sequences or DNA sequences for immunoprophylaxis.

4.3 ELISAs

ELISAs may be used in conjunction with the invention. In an ELISA assay, proteins or peptides incorporating *Borrelia* VMP-like antigenic sequences are immobilized onto a selected surface, preferably a surface exhibiting a protein affinity such as the wells of a polystyrene microtiter plate. After washing to remove incompletely adsorbed material, it is desirable to bind or coat the assay plate wells with a nonspecific protein that is known to be antigenically neutral with regard to the test antisera such as bovine serum albumin (BSA), casein or solutions of powdered milk. This allows for blocking of nonspecific adsorption sites on the immobilizing surface and thus reduces the background caused by nonspecific binding of antisera onto the surface.

After binding of antigenic material to the well, coating with a non-reactive material to reduce background, and washing to remove unbound material, the immobilizing surface is contacted with the antisera or clinical or biological extract to be tested in a manner conducive to immune complex (antigen/antibody) formation. Such conditions preferably include diluting the

antisera with diluents such as BSA, bovine gamma globulin (BGG) and phosphate buffered saline (PBS)/Tween®. These added agents also tend to assist in the reduction of nonspecific background. The layered antisera is then allowed to incubate for from about 2 to about 4 hr, at temperatures preferably on the order of about 25° to about 27°C. Following incubation, the antisera-contacted surface is washed so as to remove non-immunocomplexed material. A preferred washing procedure includes washing with a solution such as PBS/Tween®, or borate buffer.

Following formation of specific immunocomplexes between the test sample and the bound antigen, and subsequent washing, the occurrence and even amount of immunocomplex formation may be determined by subjecting same to a second antibody having specificity for the first. To provide a detecting means, the second antibody will preferably have an associated enzyme that will generate a color development upon incubating with an appropriate chromogenic substrate. Thus, for example, one will desire to contact and incubate the antisera-bound surface with a urease, alkaline phosphatase or peroxidase-conjugated anti-human IgG for a period of time and under conditions which favor the development of immunocomplex formation (*e.g.*, incubation for 2 hr at room temperature in a PBS-containing solution such as PBS/Tween®).

After incubation with the second enzyme-tagged antibody, and subsequent to washing to remove unbound material, the amount of label is quantified by incubation with a chromogenic substrate such as urea and bromocresol purple or 2,2'-azino-di-(3-ethyl-benzthiazoline)-6-sulfonic acid (ABTS) and H_2O_2, in the case of peroxidase as the enzyme label. Quantitation is then achieved by measuring the degree of color generation, *e.g.*, using a visible spectrum spectrophotometer.

4.4 Epitopic Core Sequences

The present invention is also directed to protein or peptide compositions, free from total cells and other peptides, which comprise a purified protein or peptide which incorporates an epitope that is immunologically cross-reactive with one or more anti-*Borrelia* VMP-like antibodies.

As used herein, the term "incorporating an epitope(s) that is immunologically cross-reactive with one or more anti-VMP-like antibodies" is intended to refer to a peptide or protein antigen which includes a primary, secondary or tertiary structure similar to an epitope located within a *Borrelia* VMP-like polypeptide. The level of similarity will generally be to such a degree that monoclonal or polyclonal antibodies directed against the *Borrelia* VMP-like polypeptide will also bind to, react with, or otherwise recognize, the cross-reactive peptide or protein antigen. Various immunoassay methods may be employed in conjunction with such antibodies, such as, for example, Western blotting, ELISA, RIA, and the like, all of which are known to those of skill in the art.

The identification of *Borrelia* VMP-like epitopes, and/or their functional equivalents, suitable for use in vaccines is a relatively straightforward matter. For example, one may employ the methods of Hopp, as taught in U.S. Patent 4,554,101, incorporated herein by reference, which teaches the identification and preparation of epitopes from amino acid sequences on the basis of hydrophilicity. The methods described in several other papers, and software programs based thereon, can also be used to identify epitopic core sequences (see, for example, Jameson and Wolf, 1988; Wolf *et al.*, 1988; U.S. Patent Number 4,554,101). The amino acid sequence of these "epitopic core sequences" may then be readily incorporated into peptides, either through the application of peptide synthesis or recombinant technology.

Preferred peptides for use in accordance with the present invention will generally be on the order of about 5 to about 25 amino acids in length, and more preferably about 8 to about 20 amino acids in length. It is proposed that shorter antigenic *Borrelia* VMP-like-derived peptide sequences will provide advantages in certain circumstances, for example, in the preparation of vaccines or in immunologic detection assays. Exemplary advantages include the ease of preparation and purification, the relatively low cost and improved reproducibility of production, and advantageous biodistribution.

It is proposed that particular advantages of the present invention may be realized through the preparation of synthetic peptides which include modified and/or extended epitopic/immunogenic core sequences which result in a "universal" epitopic peptide directed to *Borrelia* VMP-like and *Borrelia* VMP-like-related sequences. It is proposed that these regions

represent those which are most likely to promote T-cell or B-cell stimulation in an animal, and, hence, elicit specific antibody production in such an animal.

An epitopic core sequence, as used herein, is a relatively short stretch of amino acids that is "complementary" to, and therefore will bind, antigen binding sites on transferring-binding protein antibodies. Additionally or alternatively, an epitopic core sequence is one that will elicit antibodies that are cross-reactive with antibodies directed against the peptide compositions of the present invention. It will be understood that in the context of the present disclosure, the term "complementary" refers to amino acids or peptides that exhibit an attractive force towards each other. Thus, certain epitope core sequences of the present invention may be operationally defined in terms of their ability to compete with or perhaps displace the binding of the desired protein antigen with the corresponding protein-directed antisera.

In general, the size of the polypeptide antigen is not believed to be particularly crucial, so long as it is at least large enough to carry the identified core sequence or sequences. The smallest useful core sequence expected by the present disclosure would generally be on the order of about 5 amino acids in length, with sequences on the order of 8 or 25 being more preferred. Thus, this size will generally correspond to the smallest peptide antigens prepared in accordance with the invention. However, the size of the antigen may be larger where desired, so long as it contains a basic epitopic core sequence.

The identification of epitopic core sequences is known to those of skill in the art, for example, as described in U.S. Patent 4,554,101, incorporated herein by reference, which teaches the identification and preparation of epitopes from amino acid sequences on the basis of hydrophilicity. Moreover, numerous computer programs are available for use in predicting antigenic portions of proteins (see *e.g.*, Jameson and Wolf, 1988; Wolf *et al.*, 1988). Computerized peptide sequence analysis programs (*e.g.*, DNAStar® software, DNAStar, Inc., Madison, Wisc.) may also be useful in designing synthetic *Borrelia* VMP-like peptides and peptide analogs in accordance with the present disclosure.

Syntheses of epitopic sequences, or peptides which include an antigenic epitope within their sequence, are readily achieved using conventional synthetic techniques such as the solid phase method (*e.g.*, through the use of commercially available peptide synthesizer such as an

Applied Biosystems Model 430A Peptide Synthesizer). Peptide antigens synthesized in this manner may then be aliquoted in predetermined amounts and stored in conventional manners, such as in aqueous solutions or, even more preferably, in a powder or lyophilized state pending use.

In general, due to the relative stability of peptides, they may be readily stored in aqueous solutions for fairly long periods of time if desired, *e.g.*, up to six months or more, in virtually any aqueous solution without appreciable degradation or loss of antigenic activity. However, where extended aqueous storage is contemplated it will generally be desirable to include agents including buffers such as Tris or phosphate buffers to maintain a pH of about 7.0 to about 7.5. Moreover, it may be desirable to include agents which will inhibit microbial growth, such as sodium azide or Merthiolate. For extended storage in an aqueous state it will be desirable to store the solutions at 4°C, or more preferably, frozen. Of course, where the peptides are stored in a lyophilized or powdered state, they may be stored virtually indefinitely, *e.g.*, in metered aliquots that may be rehydrated with a predetermined amount of water (preferably distilled) or buffer prior to use.

4.5 Immunoprecipitation

The antibodies of the present invention are particularly useful for the isolation of antigens by immunoprecipitation. Immunoprecipitation involves the separation of the target antigen component from a complex mixture, and is used to discriminate or isolate minute amounts of protein. For the isolation of membrane proteins cells must be solubilized into detergent micelles. Nonionic salts are preferred, since other agents such as bile salts, precipitate at acid pH or in the presence of bivalent cations.

In an alternative embodiment the antibodies of the present invention are useful for the close juxtaposition of two antigens. This is particularly useful for increasing the localized concentration of antigens, *e.g.*, enzyme-substrate pairs.

4.6 Western Blots

The compositions of the present invention will find great use in immunoblot or western blot analysis. The anti-*Borrelia* VMP-like antibodies may be used as high-affinity primary

reagents for the identification of proteins immobilized onto a solid support matrix, such as nitrocellulose, nylon or combinations thereof. In conjunction with immunoprecipitation, followed by gel electrophoresis, these may be used as a single step reagent for use in detecting antigens against which secondary reagents used in the detection of the antigen cause an adverse background. This is especially useful when the antigens studied are immunoglobulins (precluding the use of immunoglobulins binding bacterial cell wall components), the antigens studied cross-react with the detecting agent, or they migrate at the same relative molecular weight as a cross-reacting signal.

Immunologically-based detection methods for use in conjunction with Western blotting include enzymatically-, radiolabel-, or fluorescently-tagged secondary antibodies against the toxin moiety are considered to be of particular use in this regard.

4.7 Vaccines

The present invention contemplates vaccines for use in both active and passive immunization embodiments. Immunogenic compositions, proposed to be suitable for use as a vaccine, may be prepared most readily directly from immunogenic *Borrelia* VMP-like peptides prepared in a manner disclosed herein. Preferably the antigenic material is extensively dialyzed to remove undesired small molecular weight molecules and/or lyophilized for more ready formulation into a desired vehicle.

The preparation of vaccines which contain *Borrelia* VMP-like peptide sequences as active ingredients is generally well understood in the art, as exemplified by U.S. Patents 4,608,251; 4,601,903; 4,599,231; 4,599,230; 4,596,792; and 4,578,770, all incorporated herein by reference. Typically, such vaccines are prepared as injectables. Either as liquid solutions or suspensions: solid forms suitable for solution in, or suspension in, liquid prior to injection may also be prepared. The preparation may also be emulsified. The active immunogenic ingredient is often mixed with excipients which are pharmaceutically acceptable and compatible with the active ingredient. Suitable excipients are, for example, water, saline, dextrose, glycerol, ethanol, or the like and combinations thereof. In addition, if desired, the vaccine may contain minor amounts of

auxiliary substances such as wetting or emulsifying agents, pH buffering agents, or adjuvants which enhance the effectiveness of the vaccines.

Vaccines may be conventionally administered parenterally, by injection, for example, either subcutaneously or intramuscularly. Additional formulations which are suitable for other modes of administration include suppositories and, in some cases, oral formulations. For suppositories, traditional binders and carriers may include, for example, polyalkalene glycols or triglycerides; such suppositories may be formed from mixtures containing the active ingredient in the range of about 0.5% to about 10%, preferably about 1 to about 2%. Oral formulations include such normally employed excipients as, for example, pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharine, cellulose, magnesium carbonate and the like. These compositions take the form of solutions, suspensions, tablets, pills, capsules, sustained release formulations or powders and contain about 10 to about 95% of active ingredient, preferably about 25 to about 70%.

The *Borrelia* VMP-like-derived peptides of the present invention may be formulated into the vaccine as neutral or salt forms. Pharmaceutically-acceptable salts, include the acid addition salts (formed with the free amino groups of the peptide) and those which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, oxalic, tartaric, mandelic, and the like. Salts formed with the free carboxyl groups may also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, 2-ethylamino ethanol, histidine, procaine, and the like.

The vaccines are administered in a manner compatible with the dosage formulation, and in such amount as will be therapeutically effective and immunogenic. The quantity to be administered depends on the subject to be treated, including, *e.g.*, the capacity of the individual's immune system to synthesize antibodies, and the degree of protection desired. Precise amounts of active ingredient required to be administered depend on the judgment of the practitioner. However, suitable dosage ranges are of the order of several hundred micrograms active ingredient per vaccination. Suitable regimes for initial administration and booster shots are also variable, but are typified by an initial administration followed by subsequent inoculations or other administrations.

The manner of application may be varied widely. Any of the conventional methods for administration of a vaccine are applicable. These are believed to include oral application on a solid physiologically acceptable base or in a physiologically acceptable dispersion, parenterally, by injection or the like. The dosage of the vaccine will depend on the route of administration and will vary according to the size of the host.

Various methods of achieving adjuvant effect for the vaccine includes use of agents such as aluminum hydroxide or phosphate (alum), commonly used as about 0.05 to about 0.1% solution in phosphate buffered saline, admixture with synthetic polymers of sugars (Carbopol®) used as an about 0.25% solution, aggregation of the protein in the vaccine by heat treatment with temperatures ranging between about 70° to about 101°C for a 30-second to 2-minute period, respectively. Aggregation by reactivating with pepsin treated (Fab) antibodies to albumin, mixture with bacterial cells such as *C. parvum* or endotoxins or lipopolysaccharide components of Gram-negative bacteria, emulsion in physiologically acceptable oil vehicles such as mannide mono-oleate (Aracel A) or emulsion with a 20% solution of a perfluorocarbon (Fluosol-DA®) used as a block substitute may also be employed.

In many instances, it will be desirable to have multiple administrations of the vaccine, usually not exceeding six vaccinations, more usually not exceeding four vaccinations and preferably one or more, usually at least about three vaccinations. The vaccinations will normally be at from two to twelve week intervals, more usually from three to five week intervals. Periodic boosters at intervals of 1-5 years, usually three years, will be desirable to maintain protective levels of the antibodies. The course of the immunization may be followed by assays for antibodies for the supernatant antigens. The assays may be performed by labeling with conventional labels, such as radionuclides, enzymes, fluorescents, and the like. These techniques are well known and may be found in a wide variety of patents, such as U.S. Patent Nos. 3,791,932; 4,174,384 and 3,949,064, as illustrative of these types of assays.

4.8 DNA Segments

In other embodiments, it is contemplated that certain advantages will be gained by positioning the coding DNA segment under the control of a recombinant, or heterologous,

promoter. As used herein, a recombinant or heterologous promoter is intended to refer to a promoter that is not normally associated with a DNA segment encoding a *Borrelia* VMP-like peptide in its natural environment. Such promoters may include promoters normally associated with other genes, and/or promoters isolated from any viral, prokaryotic (*e.g.*, bacterial), eukaryotic (*e.g.*, fungal, yeast, plant, or animal) cell, and particularly those of mammalian cells. Naturally, it will be important to employ a promoter that effectively directs the expression of the DNA segment in the cell type, organism, or even animal, chosen for expression. The use of promoter and cell type combinations for protein expression is generally known to those of skill in the art of molecular biology, for example, see Sambrook *et al.*, 1989. The promoters employed may be constitutive, or inducible, and can be used under the appropriate conditions to direct high level expression of the introduced DNA segment, such as is advantageous in the large-scale production of recombinant proteins or peptides. Appropriate promoter/expression systems contemplated for use in high-level expression include, but are not limited to, the *Pichia* expression vector system (Pharmacia LKB Biotechnology), a baculovirus system for expression in insect cells, or any suitable yeast or bacterial expression system.

In connection with expression embodiments to prepare recombinant proteins and peptides, it is contemplated that longer DNA segments will most often be used, with DNA segments encoding the entire peptide sequence being most preferred. However, it will be appreciated that the use of shorter DNA segments to direct the expression of *Borrelia* VMP-like peptides or epitopic core regions, such as may be used to generate anti-*Borrelia* VMP-like antibodies, also falls within the scope of the invention. DNA segments that encode *Borrelia* VMP-like peptide antigens from about 10 to about 100 amino acids in length, or more preferably, from about 20 to about 80 amino acids in length, or even more preferably, from about 30 to about 70 amino acids in length are contemplated to be particularly useful.

In addition to their use in directing the expression of *Borrelia* VMP-like peptides of the present invention, the nucleic acid sequences contemplated herein also have a variety of other uses. For example, they also have utility as probes or primers in nucleic acid hybridization embodiments. As such, it is contemplated that nucleic acid segments that comprise a sequence region that consists of at least an about 14-nucleotide long contiguous sequence that has the same sequence as, or is complementary to, an about 14-nucleotide long contiguous DNA segment of

SEQ ID NO:1 and SEQ ID NO:3 will find particular utility. Longer contiguous identical or complementary sequences, *e.g.*, those of about 20, 30, 40, 50, 100, 200, (including all intermediate lengths) and even those up to and including about 1227-bp (full-length) sequences will also be of use in certain embodiments.

The ability of such nucleic acid probes to specifically hybridize to *Borrelia* VMP-like-encoding sequences will enable them to be of use in detecting the presence of complementary sequences in a given sample. However, other uses are envisioned, including the use of the sequence information for the preparation of mutant species primers, or primers for use in preparing other genetic constructions.

Nucleic acid molecules having sequence regions consisting of contiguous nucleotide stretches of about 14, 15-20, 30, 40, 50, or even of about 100 to about 200 nucleotides or so, identical or complementary to the DNA sequence of SEQ ID NO:1 and SEQ ID NO:3, are particularly contemplated as hybridization probes for use in, *e.g.*, Southern and Northern blotting. Smaller fragments will generally find use in hybridization embodiments. wherein the length of the contiguous complementary region may be varied, such as between about 10-14 and up to about 100 nucleotides, but larger contiguous complementary stretches may be used, according to the length complementary sequences one wishes to detect.

The use of a hybridization probe of about 14 nucleotides in length allows the formation of a duplex molecule that is both stable and selective. Molecules having contiguous complementary sequences over stretches greater than 14 bases in length are generally preferred, though, in order to increase stability and selectivity of the hybrid, and thereby improve the quality and degree of specific hybrid molecules obtained. One will generally prefer to design nucleic acid molecules having gene-complementary stretches of about 15 to about 20 contiguous nucleotides, or even longer where desired.

Of course, fragments may also be obtained by other techniques such as, *e.g.*, by mechanical shearing or by restriction enzyme digestion. Small nucleic acid segments or fragments may be readily prepared by, for example. directly synthesizing the fragment by chemical means. as is commonly practiced using an automated oligonucleotide synthesizer. Also, fragments may be obtained by application of nucleic acid reproduction technology, such as PCR™, by

introducing selected sequences into recombinant vectors for recombinant production, and by other recombinant DNA techniques generally known to those of skill in the art of molecular biology.

Accordingly, the nucleotide sequences of the invention may be used for their ability to selectively form duplex molecules with complementary stretches of DNA fragments. Depending on the application envisioned, one will desire to employ varying conditions of hybridization to achieve varying degrees of selectivity of probe towards target sequence. For applications requiring high selectivity, one will typically desire to employ relatively stringent conditions to form the hybrids, *e.g.*, conditions of high stringency where one will select relatively low salt and/or high temperature conditions, such as provided by about 0.02 M to about 0.15 M NaCl at temperatures of about 50°C to about 70°C. Such selective conditions tolerate little, if any, mismatch between the probe and the template or target strand, and would be particularly suitable for isolating *Borrelia* VMP-like-encoding DNA segments. Detection of DNA segments *via* hybridization is well-known to those of skill in the art, and the teachings of U.S. Patents 4,965,188 and 5,176,995 (each incorporated herein by reference) are exemplary of the methods of hybridization analyses. Teachings such as those found in the texts of Maloy *et al.*, 1994; Segal, 1976; Prokop, 1991; and Kuby, 1994, are particularly relevant.

Of course, for some applications, for example, where one desires to prepare mutants employing a mutant primer strand hybridized to an underlying template or where one seeks to isolate *Borrelia* VMP-like-encoding sequences from related species, functional equivalents, or the like, less stringent hybridization conditions will typically be needed in order to allow formation of the heteroduplex. In these circumstances, one may desire to employ conditions such as about 0.15 M to about 0.9 M salt, at temperatures ranging from about 20°C to about 55°C. Cross-hybridizing species can thereby be readily identified as positively hybridizing signals with respect to control hybridizations. In any case, it is generally appreciated that conditions can be rendered more stringent by the addition of increasing amounts of formamide, which serves to destabilize the hybrid duplex in the same manner as increased temperature. Thus, hybridization conditions can be readily manipulated, and thus will generally be a method of choice depending on the desired results.

In certain embodiments, it will be advantageous to employ nucleic acid sequences of the present invention in combination with an appropriate means, such as a label, for determining

hybridization. A wide variety of appropriate indicator means are known in the art, including fluorescent, radioactive, enzymatic or other ligands, such as avidin/biotin, which are capable of giving a detectable signal. In preferred embodiments, one will likely desire to employ a fluorescent label or an enzyme tag, such as urease, alkaline phosphatase or peroxidase, instead of radioactive or other environmental undesirable reagents. In the case of enzyme tags, colorimetric indicator substrates are known that can be employed to provide a means visible to the human eye or spectrophotometrically, to identify specific hybridization with complementary nucleic acid-containing samples.

In general, it is envisioned that the hybridization probes described herein will be useful both as reagents in solution hybridization as well as in embodiments employing a solid phase. In embodiments involving a solid phase, the test DNA (or RNA) is adsorbed or otherwise affixed to a selected matrix or surface. This fixed, single-stranded nucleic acid is then subjected to specific hybridization with selected probes under desired conditions. The selected conditions will depend on the particular circumstances based on the particular criteria required (depending, for example, on the G+C content, type of target nucleic acid, source of nucleic acid, size of hybridization probe, *etc.*). Following washing of the hybridized surface so as to remove nonspecifically bound probe molecules, specific hybridization is detected, or even quantitated, by means of the label.

4.9 Biological Functional Equivalents

Modification and changes may be made in the structure of the peptides of the present invention and DNA segments which encode them and still obtain a functional molecule that encodes a protein or peptide with desirable characteristics. The following is a discussion based upon changing the amino acids of a protein to create an equivalent, or even an improved, second-generation molecule. The amino acid changes may be achieved by changing the codons of the DNA sequence, according to the following codon table:

TABLE 4

Amino Acids			Codons					
Alanine	Ala	A	GCA	GCC	GCG	GCU		
Cysteine	Cys	C	UGC	UGU				
Aspartic acid	Asp	D	GAC	GAU				
Glutamic acid	Glu	E	GAA	GAG				
Phenylalanine	Phe	F	UUC	UUU				
Glycine	Gly	G	GGA	GGC	GGG	GGU		
Histidine	His	H	CAC	CAU				
Isoleucine	Ile	I	AUA	AUC	AUU			
Lysine	Lys	K	AAA	AAG				
Leucine	Leu	L	UUA	UUG	CUA	CUC	CUG	CUU
Methionine	Met	M	AUG					
Asparagine	Asn	N	AAC	AAU				
Proline	Pro	P	CCA	CCC	CCG	CCU		
Glutamine	Gln	Q	CAA	CAG				
Arginine	Arg	R	AGA	AGG	CGA	CGC	CGG	CGU
Serine	Ser	S	AGC	AGU	UCA	UCC	UCG	UCU
Threonine	Thr	T	ACA	ACC	ACG	ACU		
Valine	Val	V	GUA	GUC	GUG	GUU		
Tryptophan	Trp	W	UGG					
Tyrosine	Tyr	Y	UAC	UAU				

For example, certain amino acids may be substituted for other amino acids in a protein structure without appreciable loss of interactive binding capacity with structures such as, for example, antigen-binding regions of antibodies or binding sites on substrate molecules. Since it is the interactive capacity and nature of a protein that defines that protein's biological functional activity, certain amino acid sequence substitutions can be made in a protein sequence, and, of course, its underlying DNA coding sequence, and nevertheless obtain a protein with like properties. It is thus contemplated by the inventors that various changes may be made in the

peptide sequences of the disclosed compositions, or corresponding DNA sequences which encode said peptides without appreciable loss of their biological utility or activity.

In making such changes, the hydropathic index of amino acids may be considered. The importance of the hydropathic amino acid index in conferring interactive biologic function on a protein is generally understood in the art (Kyte and Doolittle, 1982, incorporate herein by reference). It is accepted that the relative hydropathic character of the amino acid contributes to the secondary structure of the resultant protein, which in turn defines the interaction of the protein with other molecules, for example, enzymes, substrates, receptors, DNA, antibodies, antigens, and the like.

Each amino acid has been assigned a hydropathic index on the basis of their hydrophobicity and charge characteristics (Kyte and Doolittle, 1982), these are: isoleucine (+4.5); valine (+4.2); leucine (+3.8); phenylalanine (+2.8); cysteine/cystine (+2.5); methionine (+1.9); alanine (+1.8); glycine (-0.4); threonine (-0.7); serine (-0.8); tryptophan (-0.9); tyrosine (-1.3); proline (-1.6); histidine (-3.2); glutamate (-3.5); glutamine (-3.5); aspartate (-3.5); asparagine (-3.5); lysine (-3.9); and arginine (-4.5).

It is known in the art that certain amino acids may be substituted by other amino acids having a similar hydropathic index or score and still result in a protein with similar biological activity, *i.e.*, still obtain a biological functionally equivalent protein. In making such changes, the substitution of amino acids whose hydropathic indices are within ±2 is preferred, those which are within ±1 are particularly preferred, and those within ±0.5 are even more particularly preferred.

It is also understood in the art that the substitution of like amino acids can be made effectively on the basis of hydrophilicity. U.S. Patent 4,554,101, incorporated herein by reference, states that the greatest local average hydrophilicity of a protein, as governed by the hydrophilicity of its adjacent amino acids, correlates with a biological property of the protein.

As detailed in U.S. Patent 4,554,101, the following hydrophilicity values have been assigned to amino acid residues: arginine (+3.0); lysine (+3.0); aspartate (+3.0 ± 1); glutamate (+3.0 ± 1); serine (+0.3); asparagine (+0.2); glutamine (+0.2); glycine (0); threonine (-0.4); proline

(-0.5 ± 1); alanine (-0.5); histidine (-0.5); cysteine (-1.0); methionine (-1.3); valine (-1.5); leucine (-1.8); isoleucine (-1.8); tyrosine (-2.3); phenylalanine (-2.5); tryptophan (-3.4).

It is understood that an amino acid can be substituted for another having a similar hydrophilicity value and still obtain a biologically equivalent, and in particular, an immunologically equivalent protein. In such changes, the substitution of amino acids whose hydrophilicity values are within ±2 is preferred, those which are within ±1 are particularly preferred, and those within ±0.5 are even more particularly preferred.

As outlined above, amino acid substitutions are generally therefore based on the relative similarity of the amino acid side-chain substituents, for example, their hydrophobicity, hydrophilicity, charge, size, and the like. Exemplary substitutions which take various of the foregoing characteristics into consideration are well known to those of skill in the art and include: arginine and lysine; glutamate and aspartate; serine and threonine; glutamine and asparagine; and valine, leucine and isoleucine.

4.10 Site-Specific Mutagenesis

Site-specific mutagenesis is a technique useful in the preparation of individual peptides, or biologically functional equivalent proteins or peptides, through specific mutagenesis of the underlying DNA. The technique further provides a ready ability to prepare and test sequence variants, for example, incorporating one or more of the foregoing considerations, by introducing one or more nucleotide sequence changes into the DNA. Site-specific mutagenesis allows the production of mutants through the use of specific oligonucleotide sequences which encode the DNA sequence of the desired mutation, as well as a sufficient number of adjacent nucleotides, to provide a primer sequence of sufficient size and sequence complexity to form a stable duplex on both sides of the deletion junction being traversed. Typically, a primer of about 17 to 25 nucleotides in length is preferred, with about 1 to 10 residues on both sides of the junction of the sequence being altered.

In general, the technique of site-specific mutagenesis is well known in the art, as exemplified by various publications. As will be appreciated, the technique typically employs a phage vector which exists in both a single stranded and double stranded form. Typical vectors

useful in site-directed mutagenesis include vectors such as the M13 phage. These phage are readily commercially available and their use is generally well known to those skilled in the art. Double stranded plasmids are also routinely employed in site directed mutagenesis which eliminates the step of transferring the gene of interest from a plasmid to a phage.

In general, site-directed mutagenesis in accordance herewith is performed by first obtaining a single-stranded vector or melting apart of two strands of a double stranded vector which includes within its sequence a DNA sequence which encodes the desired peptide. An oligonucleotide primer bearing the desired mutated sequence is prepared, generally synthetically. This primer is then annealed with the single-stranded vector, and subjected to DNA polymerizing enzymes such as *E. coli* polymerase I Klenow fragment, in order to complete the synthesis of the mutation-bearing strand. Thus, a heteroduplex is formed wherein one strand encodes the original non-mutated sequence and the second strand bears the desired mutation. This heteroduplex vector is then used to transform appropriate cells, such as *E. coli* cells, and clones are selected which include recombinant vectors bearing the mutated sequence arrangement.

The preparation of sequence variants of the selected peptide-encoding DNA segments using site-directed mutagenesis is provided as a means of producing potentially useful species and is not meant to be limiting as there are other ways in which sequence variants of peptides and the DNA sequences encoding them may be obtained. For example, recombinant vectors encoding the desired peptide sequence may be treated with mutagenic agents, such as hydroxylamine, to obtain sequence variants.

4.11 Monoclonal Antibodies

Means for preparing and characterizing antibodies are well known in the art (See, *e.g.*, Harlow and Lane, 1988; incorporated herein by reference).

The methods for generating monoclonal antibodies (mAbs) generally begin along the same lines as those for preparing polyclonal antibodies. Briefly, a polyclonal antibody is prepared by immunizing an animal with an immunogenic composition in accordance with the present invention and collecting antisera from that immunized animal. A wide range of animal species can be used for the production of antisera. Typically the animal used for production of anti-

antisera is a rabbit, a mouse, a rat, a hamster, a guinea pig or a goat. Because of the relatively large blood volume of rabbits, a rabbit is a preferred choice for production of polyclonal antibodies.

As is well known in the art, a given composition may vary in its immunogenicity. It is often necessary therefore to boost the host immune system, as may be achieved by coupling a peptide or polypeptide immunogen to a carrier. Exemplary and preferred carriers are keyhole limpet hemocyanin (KLH) and bovine serum albumin (BSA). Other albumins such as ovalbumin, mouse serum albumin or rabbit serum albumin can also be used as carriers. Means for conjugating a polypeptide to a carrier protein are well known in the art and include glutaraldehyde. *m*-maleimidobenzoyl-N-hydroxysuccinimide ester, carbodiimide and bis-biazotized benzidine.

As is also well known in the art, the immunogenicity of a particular immunogen composition can be enhanced by the use of non-specific stimulators of the immune response, known as adjuvants. Exemplary and preferred adjuvants include complete Freund's adjuvant (a non-specific stimulator of the immune response containing killed *Mycobacterium tuberculosis*). incomplete Freund's adjuvants and aluminum hydroxide adjuvant.

The amount of immunogen composition used in the production of polyclonal antibodies varies upon the nature of the immunogen as well as the animal used for immunization. A variety of routes can be used to administer the immunogen (subcutaneous, intramuscular, intradermal, intravenous and intraperitoneal). The production of polyclonal antibodies may be monitored by sampling blood of the immunized animal at various points following immunization. A second, booster, injection may also be given. The process of boosting and titering is repeated until a suitable titer is achieved. When a desired level of immunogenicity is obtained, the immunized animal can be bled and the serum isolated and stored, and/or the animal can be used to generate mAbs.

mAbs may be readily prepared through use of well-known techniques, such as those exemplified in U.S. Patent 4,196,265, incorporated herein by reference. Typically, this technique involves immunizing a suitable animal with a selected immunogen composition. *e.g.*, a purified or partially purified LCRF protein, polypeptide or peptide. The immunizing composition is administered in a manner effective to stimulate antibody producing cells. Rodents such as mice

and rats are preferred animals, however, the use of rabbit, sheep frog cells is also possible. The use of rats may provide certain advantages (Goding, 1986), but mice are preferred, with the BALB/c mouse being most preferred as this is most routinely used and generally gives a higher percentage of stable fusions.

Following immunization, somatic cells with the potential for producing antibodies, specifically B-lymphocytes (B-cells), are selected for use in the mAb generating protocol. These cells may be obtained from biopsied spleens, tonsils or lymph nodes, or from a peripheral blood sample. Spleen cells and peripheral blood cells are preferred, the former because they are a rich source of antibody-producing cells that are in the dividing plasmablast stage, and the latter because peripheral blood is easily accessible. Often, a panel of animals will have been immunized and the spleen of animal with the highest antibody titer will be removed and the spleen lymphocytes obtained by homogenizing the spleen with a syringe. Typically, a spleen from an immunized mouse contains approximately 5×10^7 to 2×10^8 lymphocytes.

The antibody-producing B lymphocytes from the immunized animal are then fused with cells of an immortal myeloma cell, generally one of the same species as the animal that was immunized. Myeloma cell lines suited for use in hybridoma-producing fusion procedures preferably are non-antibody-producing, have high fusion efficiency, and enzyme deficiencies that render then incapable of growing in certain selective media which support the growth of only the desired fused cells (hybridomas).

Any one of a number of myeloma cells may be used, as are known to those of skill in the art (Goding, 1986; Campbell, 1984). For example, where the immunized animal is a mouse, one may use P3-X63/Ag8, X63-Ag8.653, NS1/1.Ag 4 1, Sp210-Ag14, FO, NSO/U, MPC-11, MPC11-X45-GTG 1.7 and S194/5XX0 Bul; for rats, one may use R210.RCY3, Y3-Ag 1.2.3, IR983F and 4B210; and U-266, GM1500-GRG2, LICR-LON-HMy2 and UC729-6 are all useful in connection with human cell fusions.

One preferred murine myeloma cell is the NS-1 myeloma cell line (also termed P3-NS-1-Ag4-1), which is readily available from the NIGMS Human Genetic Mutant Cell Repository by requesting cell line repository number GM3573. Another mouse myeloma cell line that may be used is the 8-azaguanine-resistant mouse murine myeloma SP2/0 non-producer cell line.

Methods for generating hybrids of antibody-producing spleen or lymph node cells and myeloma cells usually comprise mixing somatic cells with myeloma cells in a 2:1 ratio, though the ratio may vary from about 20:1 to about 1:1, respectively, in the presence of an agent or agents (chemical or electrical) that promote the fusion of cell membranes. Fusion methods using Sendai virus have been described (Kohler and Milstein, 1975; 1976), and those using polyethylene glycol (PEG), such as 37% (v/v) PEG, by Gefter *et al.*, (1977). The use of electrically induced fusion methods is also appropriate (Goding, 1986).

Fusion procedures usually produce *via*ble hybrids at low frequencies, about 1×10^{-6} to 1×10^{-8}. However, this does not pose a problem, as the *via*ble, fused hybrids are differentiated from the parental, unfused cells (particularly the unfused myeloma cells that would normally continue to divide indefinitely) by culturing in a selective medium. The selective medium is generally one that contains an agent that blocks the *de novo* synthesis of nucleotides in the tissue culture media. Exemplary and preferred agents are aminopterin, methotrexate, and azaserine. Aminopterin and methotrexate block *de novo* synthesis of both purines and pyrimidines, whereas azaserine blocks only purine synthesis. Where aminopterin or methotrexate is used, the media is supplemented with hypoxanthine and thymidine as a source of nucleotides (HAT medium). Where azaserine is used, the media is supplemented with hypoxanthine.

The preferred selection medium is HAT. Only cells capable of operating nucleotide salvage pathways are able to survive in HAT medium. The myeloma cells are defective in key enzymes of the salvage pathway, *e.g.*, hypoxanthine phosphoribosyl transferase (HPRT), and they cannot survive. The B-cells can operate this pathway, but they have a limited life span in culture and generally die within about two weeks. Therefore, the only cells that can survive in the selective media are those hybrids formed from myeloma and B-cells.

This culturing provides a population of hybridomas from which specific hybridomas are selected. Typically, selection of hybridomas is performed by culturing the cells by single-clone dilution in microtiter plates, followed by testing the individual clonal supernatants (after about two to three weeks) for the desired reactivity. The assay should be sensitive, simple and rapid, such as radioimmunoassays, enzyme immunoassays, cytotoxicity assays, plaque assays, dot immunobinding assays, and the like.

The selected hybridomas would then be serially diluted and cloned into individual antibody-producing cell lines, which clones can then be propagated indefinitely to provide mAbs. The cell lines may be exploited for mAb production in two basic ways. A sample of the hybridoma can be injected (often into the peritoneal cavity) into a histocompatible animal of the type that was used to provide the somatic and myeloma cells for the original fusion. The injected animal develops tumors secreting the specific monoclonal antibody produced by the fused cell hybrid. The body fluids of the animal, such as serum or ascites fluid, can then be tapped to provide mAbs in high concentration. The individual cell lines could also be cultured *in vitro*, where the mAbs are naturally secreted into the culture medium from which they can be readily obtained in high concentrations. mAbs produced by either means may be further purified, if desired, using filtration, centrifugation and various chromatographic methods such as HPLC or affinity chromatography.

4.12 Pharmaceutical Compositions

The pharmaceutical compositions disclosed herein may be orally administered, for example, with an inert diluent or with an assimilable edible carrier, or they may be enclosed in hard or soft shell gelatin capsule, or they may be compressed into tablets, or they may be incorporated directly with the food of the diet. For oral therapeutic administration, the active compounds may be incorporated with excipients and used in the form of ingestible tablets, buccal tables, troches, capsules, elixirs, suspensions, syrups, wafers, and the like. Such compositions and preparations should contain at least 0.1% of active compound. The percentage of the compositions and preparations may, of course, be varied and may conveniently be between about 2 to about 60% of the weight of the unit. The amount of active compounds in such therapeutically useful compositions is such that a suitable dosage will be obtained.

The tablets, troches, pills, capsules and the like may also contain the following: a binder, as gum tragacanth, acacia, cornstarch, or gelatin; excipients, such as dicalcium phosphate; a disintegrating agent, such as corn starch, potato starch, alginic acid and the like; a lubricant, such as magnesium stearate; and a sweetening agent, such as sucrose, lactose or saccharin may be added or a flavoring agent, such as peppermint, oil of wintergreen, or cherry flavoring. When the dosage unit form is a capsule, it may contain, in addition to materials of the above type, a liquid

carrier. Various other materials may be present as coatings or to otherwise modify the physical form of the dosage unit. For instance, tablets, pills, or capsules may be coated with shellac, sugar or both. A syrup of elixir may contain the active compounds sucrose as a sweetening agent methyl and propylparabens as preservatives, a dye and flavoring, such as cherry or orange flavor. Of course, any material used in preparing any dosage unit form should be pharmaceutically pure and substantially non-toxic in the amounts employed. In addition, the active compounds may be incorporated into sustained-release preparation and formulations.

The active compounds may also be administered parenterally or intraperitoneally. Solutions of the active compounds as free base or pharmacologically acceptable salts can be prepared in water suitably mixed with a surfactant, such as hydroxypropylcellulose. Dispersions can also be prepared in glycerol, liquid polyethylene glycols, and mixtures thereof and in oils. Under ordinary conditions of storage and use, these preparations contain a preservative to prevent the growth of microorganisms.

The pharmaceutical forms suitable for injectable use include sterile aqueous solutions or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersions. In all cases the form must be sterile and must be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms, such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol. polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), suitable mixtures thereof, and vegetable oils. The proper fluidity can be maintained, for example, by the use of a coating, such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. The prevention of the action of microorganisms can be brought about by various antibacterial ad antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars or sodium chloride. Prolonged absorption of the injectable compositions can be brought about by the use in the compositions of agents delaying absorption. for example, aluminum monostearate and gelatin.

Sterile injectable solutions are prepared by incorporating the active compounds in the required amount in the appropriate solvent with various of the other ingredients enumerated

above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the various sterilized active ingredients into a sterile vehicle which contains the basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum-drying and freeze-drying techniques which yield a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

As used herein, "pharmaceutically acceptable carrier" includes any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents and the like. The use of such media and agents for pharmaceutical active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active ingredient, its use in the therapeutic compositions is contemplated. Supplementary active ingredients can also be incorporated into the compositions.

The phrase "pharmaceutically acceptable" refers to molecular entities and compositions that do not produce an allergic or similar untoward reaction when administered to a human. The preparation of an aqueous composition that contains a protein as an active ingredient is well understood in the art. Typically, such compositions are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid prior to injection can also be prepared. The preparation can also be emulsified.

The composition can be formulated in a neutral or salt form. Pharmaceutically acceptable salts, include the acid addition salts (formed with the free amino groups of the protein) and which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, oxalic, tartaric, mandelic, and the like. Salts formed with the free carboxyl groups can also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, histidine, procaine and the like.

Upon formulation, solutions will be administered in a manner compatible with the dosage formulation and in such amount as is therapeutically effective. The formulations are easily administered in a variety of dosage forms such as injectable solutions, drug release capsules and the like.

For parenteral administration in an aqueous solution, for example, the solution should be suitably buffered if necessary and the liquid diluent first rendered isotonic with sufficient saline or glucose. These particular aqueous solutions are especially suitable for intravenous, intramuscular, subcutaneous and intraperitoneal administration. In this connection, sterile aqueous media which can be employed will be known to those of skill in the art in light of the present disclosure. For example, one dosage could be dissolved in 1 ml of isotonic NaCl solution and either added to 1000 ml of hypodermoclysis fluid or injected at the proposed site of infusion, (see for example, "Remington's Pharmaceutical Sciences" 15th Edition, pages 1035-1038 and 1570-1580). Some variation in dosage will necessarily occur depending on the condition of the subject being treated. The person responsible for administration will, in any event, determine the appropriate dose for the individual subject. Moreover, for human administration, preparations should meet sterility, pyrogenicity, general safety and purity standards as required by FDA Office of Biologics standards.

5.0 EXAMPLES

The following examples report the evaluation of Bb clones obtained from biopsies and blood samples from mice infected with infectious *B. burgdorferi*, an *in vivo* selection approach for detection of antigenic variation at the *vls* site and identification and characterization of the *vls* locus.

5.1 Example 1--Experimental Procedures

5.1.1 Bacterial Strains

B. burgdorferi strains B31 (ATCC 35210), Sh-2-82, and N40 were originally isolated from Ixodes scapularis ticks in the state of New York (Burgdorfer *et al.*, 1982; Schwan *et al.*, 1988b; Barthold *et al.*, 1990). These strains have been shown to be infectious in laboratory animals (Barthold *et al.*, 1990; Norris *et al.*, 1995). The high-passage B31 strain (ATCC 35210) had undergone *in vitro* passages for several years and had lost infectivity (Moody *et al.*, 1990). Nine B31 and 10 Sh-2-82 passage 5 clones had been characterized according to infectivity and described previously by Norris *et al.* (1995). An additional nine high- and

low-infectivity B31 clones were obtained from P.A. Rosa and T.G. Schwan of the Rocky Mountain Laboratories, Hamilton, MT. Infectious *B. afzelii* ACA-1 and *B. garinii* IP-90 clones were obtained by subsurface plating of organisms following isolation from experimentally infected C3H/HeN mice (A.G.B.). Spirochetes were cultured in BSK II medium as described (Norris *et al.*, 1995). The *E. coli* strains XL1-blue MRF' (Stratagene, La Jolla, CA) and BL-21 (DE3) (Novagen, Madison, WI) were used for DNA cloning and fusion protein expression, respectively.

5.1.2 Subtractive Hybridization

Subtractive hybridization was performed according to the procedure of Seal *et al.* (1992). *B. burgdorferi* total DNA was isolated from late log-phase cultures (~ 10^{10} cells) as described previously (Walker *et al.*, 1995). Total DNA of the high-passage *B. burgdorferi* B31 was subjected to ultrasonic disruption, and the resulting 0.5- to 1-kb fragments were utilized as driver DNA. The driver DNA (50 µg) was then mixed with 1 µg of total DNA from the low-passage B31 digested to completion with Sau 3 AI (target DNA). The target-driver DNA mixture was denatured and reannealled under the conditions described (Seal *et al.*, 1992). The resultant DNA mixture was ligated into *Bam*HI-digested pBluescript II SK (-) vector (Stratagene). The ligation mixture was used to transform *E. coli* XL-1 blud MRF' competent cells (Stratagene) and the transformants were plated on Luria-Bertani (LB) agar containing 100 µg/ml ampicillin, 0.5 mM isopropyl thiogalactopyranoside, and 20 µg/ml 5-bromo-4-chloro-3-indolyl-D-galactoside. LB broth cultures inoculated with white colonies were blotted to Hybond-N^+ nylon membranes (Amersham, Arlington Heights, IL) with a Bio-Dot apparatus (Bio-Rad, Hercules, CA) and screened by hybridization with [^{32}P]-labeled driver and target DNA. The clones that hybridized only to target probe but not to driver probe were partially sequenced using vector sequence-based T3 and T7 primers.

5.1.3 DNA Electrophoresis and Southern Hybridization

Total *B. burgdorferi* DNA was prepared in agarose inserts and separated in 1% Fastlane agarose gels (FMC, Rockland, ME) by pulsed-field electrophoresis as described previously (Norris *et al.*, 1995). Restriction enzyme-digested DNA fragments were separated by standard

agarose gel electrophoresis (Sambrook *et al.*, 1989). DNA bands were visualized by ethidium bromide staining. For Southern hybridization, DNA was blotted to Hybone-N^+ nylon membranes by the alkaline transfer method (Sambrook *et al.*, 1989). The blots were hybridized with [^{32}P]-labeled probes at 65°C in the presence of 1M NaCl overnight as described previously (Walker *et al.*, 1995). The blots were washed sequentially as follows: once in 2X SSC at 65°C for 15 min, twice in 1X SSC at 65°C for 15 min, and twice in 0.1X SSC at room temperature for 10 min. Autoradiography was performed using X-OMAT film (Kodak, Rochester, NY) with enhancing screens.

5.1.4 DNA Cloning and Sequence Analysis

The total plasmid DNA of B31-5A3 was prepared and treated with mung bean nuclease to open the covalently linked telomeres of the linear plasmids as described by Hinnebusch *et al.* (1990). The resulting plasmid DNA was filled in with the Klenow fragment of DNA polymerase, and an *Eco*RI linker (5'-CCGGAATTCCGG-3') was ligated onto the plasmid ends using T4 ligase. The preparation was then digested with EcoRI and ligated into EcoRI-treated λDASH II vector (Stratagene). The recombinant phages were propagated and screened by plaque hybridization with the pJRZ53 probe according to the vector manufacturer's instructions. Lambda phage DNA was purified by CsCl gradient purification method (Sambrook *et al.*, 1989).

For random cloning of the λDASH-Bb12 insert, the purified bacteriophage DNA was treated with DNase I in the presence of Mn^{++} and cloned into Eco RV-digested pBluescript II SK (-) as described previously (Demolis *et al.*, 1995). The insert DNA of λDASH-Bb12 was excised from agarose gels, purified using a Qiaex II gel extraction kit (Qiagen, Chatsworth, CA), radiolabeled, and used as a probe to screen *E. coli* XL1-blue MRF' transformants by Southern hybridization. Positive clones were sequenced as described below using T3 and T7 primers. In some instances, unsequenced regions were filled in by primer walking. The sequenced fragments were assembled using by the GELASSEMBLE program of GCG (Program Manual for the Wisconsin Package, Version 8. Genetics Computer Group, Madison.

WI). High stringency settings were applied to discriminate identical sequences from highly homologous sequences.

All plasmid and PCR™ templates were purified by Wizard columns (Promega, Madison, WI) and desalted through desalting columns (Amicon, Beverly, MA). DNA sequences were determined with an ABI377 automatic DNA sequence (Perkin-Elmer/ABI, Foster City, CA) in the DNA Core Laboratory of Department of Microbiology and Molecular Genetics at University of Texas Medical School at Houston. The GAP and PILEUP programs of GCG were used to determine sequence homology (percent similarity and identity) and to perform multiple sequence alignments, respectively. Graphical output of alignments was prepared in part through the use of the BOXSHADE program (originally programmed by K. Hofmann at Bioinformatics Group, Isrec, Switzerland and M.D. Baron at the Institute of Animal Health, Pirbright, U.K. and compiled in Pascal version for Sun Solaris/Pascal by P.A. Stockwell at University of Otago, Dunedin, New Zealand). Searches for sequence similarity were performed at the National Center for Biotechnology Information using the BLAST programs (Altschul *et al.*, 1990).

5.1.5 PCR™ Techniques

All PCR™ amplifications were performed using the thermalase PCR™ kit (Amresco, Solon, OH) in a Minicycler from MJ Research (Watertown, MA). For primer pairs containing 5'-end nested sequences F4120 (SEQ ID NO:4) and R4121(SEQ ID NO:5), a two-step program was used as follows: 96°C for 3 min, 5 cycles of denaturation at 95°C for 40 sec, annealing at 56°C for 40 sec, and extension at 72°C for 2 min, followed by 30 cycles at a higher annealing temperature at 65°C. For primer pairs without nested sequences F4064 (SEQ ID NO:6) and R4066 (SEQ ID NO:7), 35 amplification cycles of denaturation at 95°C for 40 sec, annealing at 60°C for 40 sec, and extension at 72°C for 2 min were used. The final cycles of both programs were followed by extension at 72°C for 10 min.

For RT-PCR™, total RNA was extracted from late log-phase cultures of *B. burgdorferi* B31-5A3 with a RNA purification kit (Amresco). The resulting RNA preparation was used to produce cDNA with the R4066 primer (5'-CTTTGCGAACGCAGACTCAGCA-3') (FIG. 2C).

primer R4066, and 1 µl of the RT reaction were used for PCR™ reaction as described above to produce an 198-bp fragment. The PCR™ product was then cloned into the pCR-II vector (Invitrogen, San Diego, CA) according to the supplier's manual, and the resulting clones were sequenced.

5.1.6 GST Fusion Protein Expression

A 614-bp fragment containing the *vls1* cassette was amplified by PCR™ using (+) strand primer F4120 (5'-GCGGATCCAGTACGACGGGGAAACCAG-3') and (-) strand primer R4121 (5'-GCGGATCCCCTTCTCTCTTTCTCACCATCC-3') (FIG. 2C). For cloning purposes, the inventors' added a 8-bp sequence (underlined) at the 5'-ends of both primers to create *Bam*HI sites. The resultant PCR™ products containing the entire *vls1* cassette region was cloned into the *Bam*HI site of the pGEX-2T expression vector (Pharmacia, Piscataway, NJ) to produce a GST fusion protein (Designated GST-Vls1) in *E. coli* strain BL-21(DE3) according to the supplier's instructions. The insert sequence of the recombinant plasmid was verified prior to use for protein expression. The fusion protein was purified by glutathione-Sepharose 4B column (Pharmacia) according to the manufacturer's instructions.

5.1.7 Antibodies and Immunoblotting (Western blotting)

Antisera against the GST-Vls1 fusion protein and GST as a control were prepared in rabbits by immunization of rabbits with 20 µg protein in complete Freunds adjuvant and boosting with the same amount of protein in incomplete Freunds adjuvant at 3-week intervals (Sambrook *et al.*, 1989). Nonspecific reactivity of the antiserum was removed by absorption with cell lysate of a low-infectivity B31 clone 5A2 lacking pBB28La plasmid (FIG. 1B) as described previously (Carroll and Gherardini, 1996). Antiserum against recombinant OspD was prepared in a similar manner, and monoclonal antibody H9724 reactive with *B. burgdorferi* flagellum protein (Fla) was obtained as a hybridoma culture supernatant by D.D. Thomas (University of Texas Health Science Center at San Antonio).

Late log-phase *B. burgdorferi* cultures were harvested by centrifugation and washed in phosphate-buffered-saline (PBS, 135 mM NaCl, 9 mM Na_2HPO_4, 6 mM KH_2PO_4, pH 7.2). The

organisms were resuspended at a concentration of 10^{10} cells/ml in PBS. Proteins of approximately 10^7 organisms were subjected to sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), electro-transferred to PVDF membranes (Millipore, Bedford, MA), and detected with antibody using a ECL western blot kit from Amersham according to the supplier's instructions.

5.1.8 Surface Proteolysis

Proteinase K digestion of *B. burgdorferi* B31-5A3 was performed as described previously (Norris *et al.*, 1992). Proteins of the treated organisms were separated by SDS-PAGE, electro-transferred to PVDF nitrocellulose, and reacted with antisera against GST-Vls1 or OspD or with monoclonal antibody H9724. Reactions were visualized using the ECL western blot kit.

5.1.9 Mouse Infections

The original stock of B31-5A3 (Norris *et al.*, 1995) was cultured in BSK II broth for 7 days, and the culture was diluted to a concentration of 10^6 cells/ml in BSK II broth. One hundred microliters of the diluted culture (10^5 organisms) were used to inoculate each of eight 3-week-old female C3H/HeN mice by intradermal injection at the base of the tail. Each mouse was implanted with an identification microchip for follow-up samplings during the course of infection. Four weeks after infection, the organisms were isolated by inoculating 50 µl of blood or a full-thickness biopsy (~2 mm in diameter) of the ear into 6 ml of BSK II broth. All 5 ear cultures and 6 of 8 blood cultures were positive. Clonal populations of *B. burgdorferi* isolates from C3H/HeN mice were obtained by subsurface plating (Norris *et al.*, 1995). The first passages of these cultures were frozen in BSK II medium with 20% glycerol at -70°C as stocks for further study. The *vls* cassette region at the expression site was amplified by PCR™ using primers F4120 and R4066 (FIG. 2C) and sequenced using the same set of primers. Samples of the frozen stocks (~ 3 µl) were scraped from the surface, thawed, and added directly into PCR™ tubes as the DNA template source to minimize possible variation during *in vitro* cultivation.

Serum sample were also collected from each mouse before infection and 4 weeks after initial infection and stored at -70°C for immunoblot analysis.

5.2 Example 2--Antigenic Variation at *vls* Site

Bb clones isolated from ear biopsies and blood samples obtained 4 weeks post inoculation of C3H/HEN mice with the infectious B31 clone 5A3 were evaluated. A flow diagram is provided in FIG. 8. Ear punch biopsies ~ 2mm in diameter were obtained from 5 of 8 mice. The cultures were named according to their source (*i.e.* m1e4 refers to mouse 1, ear culture, 4 weeks). Clonal populations were obtained by plating on passage of culture BSKY agar plates, and 12 colonies of each isolate were selected, cultured briefly in 2ml BSKY medium, and frozen. Individual clones were designated m1e4A, m1e4B, *etc.* Samples of these clones were subjected to amplification of the *vls* cassette present in the *vls*E expression site by using primers in the "constant" regions on either side of the cassette. The resulting PCR™ products were sequenced directly.

Surprisingly, antigenic variation did not occur by substitution of the entire *vls* cassette at the expression site (*e.g. vls1*) with a single, intact "silent" *vls* cassette (*e.g. vls2* through *vls16*) (see FIG. 9). Examination of 5 clones from the ear of mouse 1 (FIG. 10) indicated that the *vlsE* sequences of each clone differed from 1) the original sequence (*vls1*); 2) one another; and, 3) each of the silent *vls* cassettes. These results were verified by examination of one clone from each ear or blood isolate from the eight mice (FIG. 11). Each one of the mouse isolate sequences appeared to be comprised of a mosaic of segments from several different silent *vls* cassettes (between 7 and 11 in preliminary analyses). Sequence variability was restricted to the *vls* cassette region delimited by the ' sequences, and in all cases the open reading frame was preserved. The *vlsE* cassette regions of controls consisting of five clonal populations from the inoculating culture were identical to the original *vls1* sequence, as were the sequences obtained from cultures passed 2 to 4 times *in vitro* (one week per passage). Therefore the rearrangement process appears to be activated *in vivo*, and does not occur at a rapid rate *in vitro*.

5.3 Example 3--Identification of the 28-kb linear plasmid, pBB28La

B. burgdorferi strains generally exhibit loss of infectivity following 10 to 17 *in vitro* passages (Johnson *et al.*, 1984; Schwan *et al.*, 1988a; Norris *et al.*, 1995), coinciding with the loss of plasmids (Barbour, 1988; Xu *et al.*, 1996). It was hypothesized that the decreased infectivity occurring during *in vitro* passage of Lyme disease *Borreliae* is due to loss of genetic content, specifically plasmids encoding virulence factors. Therefore, the inventors' expected to identify some of these virulence factors by directly comparing the plasmid content of the organisms differing in infectivity.

One of the complications involved in studying *B. burgdorferi* plasmids is that many plasmids are in the 20 kb to 40 kb size range (Xu and Johnson, 1995), making it difficult to resolve plasmids with similar sizes by standard electrophoretic techniques. In addition, mutagenic techniques and other genetic manipulation tools are in an early stage of development in *B. burgdorferi* (Samuels *et al.*, 1994; Rosa *et al.*, 1996), thereby limiting the ability to examine the importance of these plasmids in pathogenesis by direct genetic approaches.

To overcome these limitations, the inventors' utilized a simple subtractive hybridization technique to enrich and eventually identify sequences present only in high-infectivity organisms.

Total DNA from a high-infectivity (low-passage) B31 strain was digested to completion with Sau 3AI (target DNA). The target DNA was mixed with a 50-fold excess of total DNA from a low-infectivity (high passage) B31 derivative that had been sheared by ultrasonication (driver DNA). The DNA mixture was denatured and allowed to reanneal. DNA fragments in the resultant DNA preparation in which the Sau 3 AI sites were regenerated were ligated into the *Bam*HI site of pBluescript SK (-). A total of 63 clones were isolated and screened by Southern hybridization using the target DNA and driver DNA as probes, respectively. Eight of these clones hybridized with target DNA but not to driver DNA.

The inserts of the eight clones were partially sequenced using the vector-based primers, and the sequences were subjected to database searches for sequence similarity. One of the clones, designated pJRZ53, contained a 562-base pair (bp) Sau 3 AI fragment with a single,

contiguous open reading frame. The predicted amino acid sequence of this open reading frame was compared to Vmps of *B. hermsii*, and showed 27.2% identity and 56.8% similarity to Vmp17. Based on this sequence similarity, the pJRZ53 insert was called a *vmp*-like sequence (*vls*). The pJRZ53 insert exhibited a lower degree of amino acid sequence similarity with *B. burgdorferi* B31 outer surface protein C (OspC) (26.6% identity and 44.5% similarity).

To identify the genomic location of the *vls* sequence, the pJRZ53 insert was hybridized with Southern blots of total B31 DNA separated by pulsed-field electrophoresis. A DNA band migrating at 28 kb hybridized to the probe (see FIG. 1B) and was determined to be a linear plasmid by two-dimensional agarose gel electrophoresis and restriction mapping. This *vls*-containing plasmid of *B. burgdorferi* B31 was designated pBB28La.

5.4 Example 4--Correlation between pBB28La and infectivity

This example illustrates the determination whether or not the *vls*-encoded function was required for infection. Previous studies (Norris *et al.*, 1995), had indicated that clones of low-passage *B. burgdorferi* strains B31 and Sh-2-82 exhibited two distinct infective phenotypes when tested in C3H/HeN mice. A representative high-infective clone had a median infectious dose (ID_{50}) of 1.8×10^2 organisms, whereas the low-infective clone tested showed a much higher ID_{50} ($\geq 1 \times 10^5$ organisms).

It was reasoned that if the *vls*-encoded function is important for virulence, all clones with the high-infective phenotype should contain the *vls*-containing plasmid, and loss of this plasmid would result in low-infective phenotype. To test this hypothesis, the pJRZ53 probe was hybridized with total DNA from both high- and low-infective *B. burgdorferi* clones. All nine B31 clones tested had a plasmid banding pattern almost identical to each other when visualized by ethidium bromide staining (FIG. 1A).

However, hybridization of pJRZ53 with the blot made from the same gel revealed that all five high-infective B31 clones possessed the *vls*-containing pBB28La plasmid, whereas only 1 or 4 low-infective clones (B31-5A10) had this plasmid (FIG. 1B). It appears that the low-infective B31-5A10 clone is lacking another plasmid that correlates with infection. Nine additional low-passage B31 clones obtained from P.A. Rosa and T.G. Schwan (Rocky

Mountain Laboratories, Hamilton, MT) exhibited a similar pattern; all six of the high-infective clones contained pBB28La, whereas only 1 of 3 low-infective clones contained the plasmid, based on hybridization with the pJRZ53 probe. These results indicated a strong correlation of pBB28La with the high-infective phenotype in clonal populations of *B. burgdorferi* B31.

To verify the correlation found in strain B31, ten previously characterized clones of strain Sh-2-82 (Norris *et al.*, 1995) were examined. A pBB28La homolog was detected in seven high-infective clones but not in three low-infective clones of strain Sh-2-82. Similar studies revealed the presence of a single *vls*-containing plasmid in infectious *B. afzelii* ACA-1 and *B. garinii* IP-90 strains. In contrast to the multiple *vmp*-containing linear plasmids in *B. hermsii*, only one *vls*-containing plasmid was found in each of the Lyme disease isolates tested under the hybridization conditions employed. These *vls*-containing plasmids migrated consistently at ~ 28 kb in agarose gels in the strains examined.

Out of a total of 32 clones or strains examined, all 22 clones or strains with the high-infective phenotype contained pBB28La, and only 2 of 10 low-infective clones possessed this plasmid (Table 2). The Southern hybridization studies indicated that the *vls*-containing plasmid is associated with infectivity and therefore may encode essential virulence factor(s).

5.5 Example 5--Cloning and Sequencing of the *vls* locus

A particular clonal population of *B. burgdorferi* B31 (clone B31-5A3) was utilized in order to minimize clonal variation. B31-5A3 has a high-infective phenotype (Norris *et al.*, 1995) and possesses the pBB28La plasmid (FIG. 1B, lane 3). pJRZ53 was shown to hybridize with a single 14-kb fragment generated by digestion of B31-5A3 plasmid DNA with *Eco*RI. However, treatment of the B31-5A3 plasmid DNA with *Pst*I, Sau 3 AI, or *Rsa*I each yielded multiple fragments ranging in size from ~ 400 bp to 4,000 bp that hybridized with the probe, denoting the presence of multiple copies of the *vls* sequence.

The 14-kb *Eco*RI fragment was cloned into λDASHII to permit a detailed analysis of this region. The 14-kb fragment was predicted to have a covalently-closed telomere at one end. Therefore, a technique developed by Hinnebusch *et al.* (1992) was utilized to open the telomeric loop with mung bean nuclease and attach an *Eco*RI linker, thereby permitting ligation into the

cloning vector. A lambda clone, designated λDASH-Bb12, was isolated that contained the 14-kb *B. burgdorferi* DNA fragment, as confirmed by restriction and hybridization. An internal *Eco*RI site was found to divide the λDASH-Bb12 insert into two smaller 4- and 10-kb fragments; an independently-derived clone containing the 10-kb fragment was also isolated during screening of the library. To verify that the 4-kb and 10-kb *Eco*RI fragments were physically linked in the native *B. burgdorferi* plasmid, the region containing the internal *Eco*RI site was amplified using *B. burgdorferi* B31-5A3 DNA as the template. The resulting PCR™ product had a sequence identical to that of the corresponding region of λDASH-Bb12, indicating that the 4- and 10-kb *Eco*RI fragments are contiguous in pBB28La. Restriction digestion of *B. burgdorferi* plasmid DNA at this *Eco*RI site was not efficient, whereas complete cleavage was obtained consistently at the same site in the λDASH construct. Similar incomplete digestion has been observed with certain restriction sites in *B. burgdorferi* chromosomal DNA (Casjens *et al.*, 1995) and may be related to DNA modification (Hughes and Johnson, 1990).

A random cloning, "shotgun" strategy was utilized to sequence nearly 10 kb of the λDASH-Bb12 insert. A total of 80 random clones were sequenced using vector-based primers. Additional sequencing reactions were carried out to fill the gaps between the sequenced regions by primer-walking. The resulting assembled sequences had an average of 5-fold coverage. A short segment (~ 200 bp) 1227 bp from the right telomeric end has been refractory to sequencing by a number of techniques. In contrast to the overall low guanosine-cytosine (G+C) content of the *B. burgdorferi* genome (~ 28%), the *vls* locus has a G+C content of 50%.

5.6 Example 6--Organization of the *vls* Locus

The sequence data revealed an extensive *vls* locus within the 10-kb *Eco*RI fragment consisting of an expression site (designated *vlsE*) and 15 *vls* cassettes that are highly homologous to the central portion of *vlsE* (SEQ ID NO:1 and SEQ ID NO:3). The presence of the *Eco*RI linker sequence between the insert DNA and the vector sequence defined the location of the right telomeric end. *VlsE* is located 82 bp from the right telomere of pBB28La. It possesses two unique sequences at each of the 5' and 3' regions and a 570-bp *vls* cassette in the middle which was designated as the *vls1* cassette (FIG. 2B). The *vls1* cassette is flanked at

either end by the 17-bp direct repeat sequence 5'- GAGGGGGCTATTAAGGA -3' (SEQ ID NO:8) encoding the amino acids EGAIK. An array of 15 *vls* cassettes begins approximately 500 bp upstream of *vlsE* on the same plasmid (FIG. 2A). The *vls1* cassette and the other *vls* cassettes (*vls2* through *vls16*) share 90.0% to 96.1% nucleotide sequence identity and 76.9% to 91.4% predicted amino acid sequence identity. The 17-bp direct repeat is conserved in nearly all of the upstream *vls* cassette sequences.

The *vlsE* gene of *B. burgdorferi* B31-5A3 is predicted to encode a 356 amino acid protein with a molecular mass of 35,986 daltons (FIG. 2C). A consensus ribosome binding site and consensus -35 and -10 sigma -70-like promoter sequences are located upstream of the predicted translational start site. VlsE contains a putative lipoprotein leader sequence with an apparent signal peptidase II cleavage site (FINC) (Wu and Tokunaga, 1986) which resembles those of other *Borrelial* lipoproteins, including OspC (Fuchs *et al.*, 1992). Cleavage of the 18 amino acid signal peptide would result in a mature polypeptide with a calculated molecular mass of 33,956 daltons and an isoelectric point (pI) of 7.3. Except for the putative leader peptide, VlsE is predominantly hydrophilic.

VlsE shows 37.4% identity and 57.8% similarity homology at the amino acid level and 58.8% identity at the nucleotide level to *vmp17* of *B. hermsii* (FIG. 3A). *VlsE* shares a lower level of homology to *B. burgdorferi ospC* at both the nucleotide (41.6% identity) and amino acid (26.3% identity and 47.5% similarity) levels. The particular *vlsE* allele contained in *B. burgdorferi* B31 clone 5A3 has been designated *vlsE1*, to distinguish it from variant *vlsE* alleles.

An additional 15 *vls* cassettes (474 to 594 bp in length) were identified ~500 bp upstream of *vlsE* (FIG. 2A and FIG. 3B, SEQ ID NO:1 and SEQ ID NO:3). These cassettes are oriented in the opposite direction to *vlsE* and are arranged in a head-to-tail fashion in a nearly contiguous open reading frame interrupted only by a stop codon in cassette *vls11* and two frame shifts in cassettes *vls14* and *vls16*. None of these *vls* cassettes have recognizable ribosome binding sites or promoter sequences; therefore they are thought to be nonexpressed or 'silent'. The ends of the *vls* cassettes were defined by alignment with the *vls1* cassette (FIG. 3B). In general, the *vls* cassettes have the same 17-bp direct repeat at either end; one exception is the joint region between *vls9* and *vls10*, where only 10 identical nucleotides were identified. The

first *vls* cassette (*vls2*) lacks the first 126 bp of the *vls* cassette sequence, but contains a 55-bp sequence which is identical to the 5′ region of *vlsE*, coding for the last 11 amino acids of the leader peptide and the first 7 amino acids of the putative mature VlsE. The *vls7* cassette contains a 105-bp deletion relative to *vls1* in the 5′ region. The *vls8* and *vls10* cassettes are lacking the first 54 nucleotides of the cassette. The last cassette in this array, *vls16*, is truncated at the 3′ end and is followed by an apparent noncoding region. The 562-bp insert of pJRZ53 was localized to the joining region between vls8 and vls9 by sequence comparison.

The *vls* cassettes contain six highly conserved regions which are interspersed by six variable regions (VR) at both the nucleotide and amino acid levels. FIG. 3B shows an alignment of the predicted amino acid sequences for all 16 *vls* cassettes identified. Except for occasional codon changes and the deletions mentioned previously, the conserved regions are almost identical in all cassettes. On the other hand, the *vls* cassettes are distinguished from each other by considerable sequence variations limited predominantly to the six variable regions (VR-I through VR-VI). The variable regions range from 21 bp (VR-VI) to 63 bp (VR-IV) in length. With exception of an insertion of a TAG stop codon in *vls11* and TG insertions in *vls14* and *vls16* resulting in frameshifts, all deletions and insertions are nucleotide triplets, indicating preservation of the open frame. The sequence variations at most polymorphic positions result in conservative amino acid changes, suggesting that certain amino acids are required at these positions for function. Even within the six variable regions, there is a clear sequence conservation. For example, the variable sequences in VR-I are interspersed by stretches of identical sequences ranging from 6 to 9 bp, as reflected in the predicted amino acid sequences (FIG. 3B).

5.7 Example 7--Expression of *vlsE*

To determine whether *vlsE* is transcribed, reverse transcriptase-polymerase chain reaction (RT-PCR™) was utilized to amplify a 3′ region of *vlsE* (191bp) from total RNA of *in vitro* cultured B31-5A3. After the reverse transcriptase reaction, PCR™ amplification, and agarose electrophoresis, a band of the expected size was observed by ethidium bromide staining. The RT-PCR™ product was cloned into the pCRII vector and the recombinant plasmids were sequenced. Three independently derived recombinant plasmids contained DNA

sequences identical to the corresponding region of *vlsE*, demonstrating that *vlsE* is transcribed *in vitro*. No RT-PCR™ products were observed in the agarose gel if reverse transcriptase was omitted from the reaction, confirming that the RT-PCR™ products were derived from the mRNA of B31-5A3.

To identify the protein product of *vlsE*, an internal 614-bp fragment containing *vls1* was amplified by polymerase chain reaction (PCR™) and cloned into the pGEX-2T expression vector to produce a glutathione-S-transferase (GST)-Vls1 fusion protein in *E. coli*. Rabbit antiserum against the GST-Vls1 fusion protein was used to probe protein blots of *B. burgdorferi* B31-5A2 and B31-5A3 clones. The low-infectivity clone B31-5A2 was used as a negative control for immunoblot analysis, because it lacks pBB28La (FIG. 1B, lane 2). The antiserum detected a protein with an M_r of approximately 45,000 daltons in the high-infectivity clone B31-5A3 but not in the low-infectivity clone B31-5A2 (FIG. 6, lanes 10 and 11). Neither the preimmune serum nor antiserum against GST alone reacted with this protein. The size of the protein identified by immunoblot analysis is larger than the predicted molecular mass of 33,956 daltons. Attachment of a lipid moiety to the N-terminus of VlsE by signal peptidase II may contribute to the altered electrophoretic mobility.

5.8 Example 8--Surface Localization of VlsE

The presence of a putative lipoprotein leader peptide and the overall hydrophilic nature of VlsE raised the possibility that it is attached to the bacterial membrane *via* a lipid anchor. To test this possibility, *B. burgdorferi* B31-5A3 was incubated in the presence of [^{3}H]-palmitate as described previously (Norris *et al.*, 1992). VlsE was radiolabelled by [^{3}H]-palamitate along with other *B. burgdorferi* lipoproteins, suggesting that VlsE is a lipoprotein.

Exposure of viable *B. burgdorferi* 5A3 to proteinase K produced results consistent with the surface localization of VlsE. VlsE was degraded by proteinase K in as little as 10 min (FIG. 4A), even though the organisms appeared intact by dark-field microscopy. Consistent with previous study (Norris *et al.*, 1992), *B. burgdorferi* OspD protein was also removed by proteinase K treatment (FIG. 4B). In contrast, the Fla subunit of the periplasmic flagella was

not affected by proteinase K (FIG. 4C), providing further evidence that the outer membranes of the organisms remained intact during the proteinase K treatment.

5.9 Example 9--Genetic Variation at the *vlsE* Site

The similarity of the *vls* locus to the *vmp* system of *B. hermsii* prompted the question whether genetic recombination between the expressed and silent *vls* cassettes could be demonstrated in the mammalian host. The overall experimental design is illustrated in FIG. 5A. *B. burgdorferi* B31-5A3, inoculated directly from a frozen stock, was cultured for seven days and used to intradermally inject a group of eight female C3H/HeN mice (10^5 organisms per mouse). *B. burgdorferi* was re-isolated four weeks after the initial infection. To retain the infected mice for multiple samplings at different periods of infection, only ear punch biopsy and blood specimens were taken to culture the organisms. A total of five ear and six blood isolates were examined. To examine possible genetic heterogeneity within the mouse isolates, 16 *B. burgdorferi* clones of each isolate were obtained by colony formation on agarose plates and preserved by freezing. One clone from each of the isolates was used as a source of template DNA to amplify the expressed *vls* cassette sequence using primers F4120 and R4066 specific for the 5' and 3' unique regions of *vlsE*, respectively (FIG. 2C). The first passage frozen stock was used to provide DNA template for PCR™ amplification to avoid possible variation during *in vitro* culture. The PCR™ products were sequenced directly using the same set of primers. The *B. burgdorferi* clones and associated sequences derived from the 4-week isolates were designated by a combination of mouse number (m1 to m8), tissue source (e for ear and b for blood), week post infection (4), and a clone designation (A to P) for the 16 clones from each isolate.

When compared with the parental *vlsE* of the clone B31-5A3 (allele *vlsE1*) inoculation, multiple base substitutions, deletions and insertions were found within the *vls* cassette region of *vlsE*, making each allele unique. These changes resulted in numerous differences in the predicted amino acid sequences (FIG. 5B). As found in the silent *vls* cassettes (FIG. 3B), these mutations were primarily confined within the six variable regions. The variable sequences at almost all positions in the 11 *vlsE* alleles could be found in the corresponding regions of the silent *vls* cassettes. For example, the *m1e4A* and *m5e4A* alleles have VR-I and VR-II identical

to *vls4*, whereas the VR-I and VR-III regions of *m6b4A* are identical to the same regions of *vls10* (FIG. 5B). These results indicated that changes in the original *vls1* cassette have originated from the silent *vls* cassettes *via* genetic recombination. In contrast, the sequences on either side of the *vls* cassette remained unchanged in the 11 alleles examined (FIG. 5B).

Based on the gene conversion mechanism in *vmp* systems, it was initially hypothesized that if genetic recombination occurred at the *vlsE* site, the expressed *vls* cassette (*vls1* in this case) would be replaced completely by a single silent cassette flanked by the 17-bp direct repeat. However, careful examination revealed that none of the 11 *vlsE* alleles examined were identical to any of the silent *vls* cassettes identified to date. Rather, each allele appeared to be a mosaic of segments from several different silent *vls* cassettes. For instance, although *mle4A* shares common sequence with *vls4* throughout VR-I and VR-II, its VR-III and VR-VI are the same as *vls 10* and *vls2*, respectively. Interestingly, the VR-IV and VR-V regions of *mle4A* appear to be hybrids of portions of *vls10* and *vls5* and *vls3* and *vls5*, respectively. Similar patterns can also be found in the rest of these *vlsE* alleles. These observations suggested that segments, but not entire regions, of the silent *vls* cassettes were recombined into the expression site. Comparison to the silent cassette sequences at the nucleotide level suggested that 6 to 11 separate recombination events occurred in each of the clones isolated from mice 4 weeks post inoculation. This type of combinatorial reactions could potentially result in millions of different *vlsE* alleles.

To determine whether the clonal populations from a single mouse also exhibited similar sequence variations, four additional clones of the blood and ear isolates from mouse 1 were chosen to determine the DNA sequence at the *vlsE* site. The five clones (m1b4A-E) of the blood isolate had sequences identical to each other, although they showed considerable sequence differences from the parental *vlsE* as represented by *m1b4A* (FIG. 5B). In contrast, the sequences from the five clones of the ear isolate differed substantially both from the parental *vls1* cassette and from each other (FIG. 5C). Consistent with the 11 *vlsE* alleles from different isolates, the sequence variations from the same ear isolate were also concentrated in six variable regions (FIG. 5C). Each of these clones again contained a unique combination of sequences identical to portions of several silent *vls* cassettes. For example, *mle4C* contained VR-I of *vls12*, VR-II of *vls4*, VR-III of *vls8*, and VR-IV and VR-VI of *vls11*. The homogeneous nature

of *B. burgdorferi* clones derived from the blood isolate of mouse 1 may be due to the presence of relatively few organisms in the blood as compared to ear biopsies, resulting essentially in cloning by limiting dilution. Alternatively, selection imposed by the host immune response in different tissue environments may affect diversity of *vlsE* variants.

The sequence variations in the clonal populations of the mouse isolates may also arise from background heterogeneity of the stock culture of the clone B31-5A3 occurring during *in vitro* culture, because the original clone was cultured 7 days prior to the inoculation of C3H/HeN mice. To test this possibility, the stock culture of B31-5A3 was inoculated into BSK II medium and cultured sequentially in two *in vitro* passages of 7 days (14 days total). PCR™ products amplified from the *vlsE* cassette region were obtained using a sample of this culture as template and either sequenced directly or cloned into the PCR™ II vector and sequenced. Two sets of PCR™ products and four independently derived recombinant plasmids containing the PCR™ products all had sequences identical to the initial *vlsE* sequence. These results indicated that the sequence variations did not occur at high frequency at the *vlsE* site prior to the inoculation of mice.

5.10 Example 10--Changes in Antigenicity of the VlsE Variants

The promiscuous genetic recombination at the *vlsE* site and the putative surface location of VlsE suggested that sequence variations in the *vlsE* alleles result in changes in antigenicity. Nine clonal populations carrying unique *vlsE* alleles (see FIG. 5B) were subjected to immunoblot analysis. Although a similar amount of total proteins were loaded into the gel as indicated by reactivity to antibody against *B. burgdorferi* flagellin protein (FIG. 6A), these VlsE variants exhibited a dramatic decrease in reactivity to the antiserum against the GST-Vls1 fusion protein (FIG. 6B). The mouse isolates containing *m1b4A* and *m3b4A* alleles had bands which were weakly reactive with the antiserum (FIG. 6, lanes 2 and 5). The other clones examined exhibited faint bands that were visible only with a longer chemiluminescent exposures of the membrane. These reactive bands migrated at lower M_rs than VlsE expressed by the parental clone B31-5A3. indicative of changes in either size or conformation. No reactive bands were observed in clone B31-5A2, which lacks the pBB28L.a plasmid. The decreased reactivity of mouse isolates with antiserum against the parental Vls1 cassette region

indicated that the sequence differences in these VlsE variants (FIG. 5B) resulted in changes in important cassette region epitopes and hence antigenic variation.

5.11 Example 11--*In Vivo* Expression of Vls E and Induction of Antibodies in Infected Humans and Animals

Sera from the mice in experiments outlined in FIG. 5A were tested for reactivity with VlsE as a means of determining whether this protein is expressed *in vivo*. Serum obtained from mouse 1 on 28 days post inoculation with B31-5A3 was reacted with immunoblots of 5A3 (expressing VlsE), 5A2 (lacking vlsE), the GST-Vls1 fusion protein, GST as a control, and two clones isolated from mouse 1 on day 28 (M1e4A and M1b4A). The results shown in FIG 6C indicated that the C3H/HeN mice infected with *B. burgdorferi* mounted a strong antibody response to VlsE. Although the prebleed serum of mouse 1 had no detectable reactivity, the serum sample collected from the same mouse 4 weeks after initial infection with *B. burgdorferi* B31-5A3 reacted strongly with the GST-Vls1 fusion but not with GST alone, indicating expression of VlsE in the mammalian host. The same serum also had a strong reactivity with the VlsE protein of *B. burgdorferi* B31-5A3, whereas no detectable VlsE band was observed with *B. burgdorferi* B31-5A2. In contrast, the VlsE variant M1e4A exhibited decreased reactivity when reacted with the same mouse serum as shown in FIG. 6C.

Since the C3H/HeN mice were infected with a large number (10^5) of the organisms (see FIG. 5A), it was possible that the antibody response against VlsE had resulted from the initial inoculum. To test this possibility, sera from white-footed mice (Peromyscus leucopus) infected with *B. burgdorferi* B31 via tick bite and from human Lyme disease patients were used to react with the similar immunoblots. The representative results depicted in FIG. 6D showed that tick-infected Peromyscus mice also had strong reactivity to the VlsE protein of *B. burgdorferi* B31-5A3 and GST-Vls1 fusion protein but not with GST alone. These results were further confirmed with sera from Lyme disease patients (FIG. 6E). A representative serum sample from a clinically diagnosed patient with early Lyme disease symptoms contained highly reactive antibody against the VlsE protein of B31-5A3 and GST-Vls1 fusion protein (FIG. 6E). Similar to the serum from the C3H/HeN mouse (FIG. 6C), the sera from the Peromyscus leucopcus mouse (FIG. 6D) and the Lyme disease patient (FIG. 6E) had little reactivity to the

VlsE variant M1e4A. These results indicate that VlsE is expressed and is highly immunogenic in the mammalian host, but that genetic variation may generate unique VlsE variants which are no longer fully recognized by the immune response against the parental VlsE. They also indicate that antibodies generated against VlsE may be useful in immunodiagnosis of Lyme disease.

These results indicated that VlsE is expressed and is highly immunogenic in the mammalian host, but that genetic variation can generate unique VlsE variants which are no longer fully recognized by the immune response against the parental VlsE. Additional experiments have shown that some sera from Lyme disease patients also have reactivity with the GST-Vls1 fusion protein and VlsE of *B. burgdorferi B31-5A3*, but not with some of the VlsE variants, thus further supporting the expression and antigenic variation of VlsE *in vivo*.

TABLE 5

Correlation of pBB28La with Infectivity

Strain	Strains containing pBB28La/total strains tested	
	High-infectivity phenotype	Low-infectivity phenotype
B. burgdorferi B31	12/12	2/7
B. burgdorferi Sh2-82	7/7	0/3
B. burgdorferi N40	1/1	ND[a]
B. afzelii ACA-1	1/1	ND
B. garinii IP-90	1/1	ND
Total	22/22	2/10

[a]Not determined

6.0 REFERENCES

The following literature citations as well as those cited above are incorporated in pertinent part by reference herein for the reasons cited in the above text.

5,436,000 Flagella-less *Borrelia*

5,434,077 *Borrelia burgdorferi* strain 257

5,403,718 Methods and antibodies for the immune capture and detection of *Borrelia burgdorferi*

5,385,826 Diagnostic assay for Lyme disease

5,324,630 Methods and compositions for diagnosing Lyme disease

5,283,175 Genus-specific oligomers of *Borrelia* and methods of using same

5,279,938 Sensitive diagnostic test for Lyme disease

5,246,844 Virulence associated proteins in *Borrelia burgdorferi*

5,217,872 Method for detection of *Borrelia burgdorferi* antigens

5,187,065 Method and materials for detecting Lyme disease

5,178,859 Vaccine against Lyme disease

5,155,022 Assay for Lyme disease

Altschul, Gish, Miller, Myers, Lipman, "Basic local alignment search tool," *J. Mol. Biol.*, 215:403-410, 1990.

Balmelii and Piffatetti, "Analysis of the genetic polymorphism of *Borrelia burgdorferi* sensu lato by multilocus enzyme electrophoresis," *Int. J. Syst. Bacteriol.*, 46:167-172, 1996.

Barbour, "Plasmid analysis of *Borrelia burgdorferi*, the Lyme disease agent," *J. Clin. Microbiol.*, 42:475-478, 1988.

Barbour, "Plasmid analysis of *Borrelia burgdorferi*, the Lyme disease agent," *J. Clin. Microbiol.*, 26:475-478, 1988.

Barbour, "Linear DNA of *Borrelia* species and antigenic variation," *Trends Microbiol.*, 1:236-239, 1993.

Barbour and Garon, "Linear plasmids of the bacterium *Borrelia* burdorferi have covalently closed ends," *Science*, 237:409-411, 1987.

Barbour, Burman, Carter, Kitten, Bergstrom, "Variable antigen genes of the relapsing fever agent *Borrelia* hermsii are activated by promoter addition," *Mol. Microbiol.*, 5:489-493, 1991a.

Barbour, Carter, Burman, Freitag, Garon, Bergstrom, "Tandem insertion sequence-like elements define the expression site for variable antigen genes of *Borrelia* hermsii," *Infect. Immun.*, 59:390-397, 1991b.

Barbour *et al.*, "Structural analysis of the variable major proteins of *Borrelia hermsii*," *J. Exp. Med.*, 158:2127-2140, 1983.

Barbour *et al.*, "Variable major proteins of *Borrelia hermsii*," *J. Exp. Med.*, 156:1312-1324, 1982.

Barstad *et al.*, "Variable major proteins of *Borrelia hermsii*. Epitope mapping and partial sequence analysis of CNBr peptides," *J. Exp. Med.*, 161:1302-1314, 1985.

Barthold, "Antigenic stability of *Borrelia burgdorferi* during chronic infections of immunocompetent mice," *Infect. Immun.*, 61:4955-4961, 1993.

Barthold, Moody, Beck, "Suspectibility of laboratory rats to isolates of *Borrelia burgdorferi* from different geographic areas," *Am. J. Trop. Med. Hyg.*, 42:596-600, 1990.

Borst and Geaves, "Programmed gene rearrangements altering gene expression," *Science*, 235:658-667, 1987.

Borst, Bitter, McCulloch, Leeuwen, Rudenko, "Antigenic variation in malaria," *Cell*, 82:104, 1995.

Burgdorfer, Barbour, Hayes, Benach, Grunwaldt, Davis, "Lyme disease, a tick-borne spirochetosis?," *Science*, 216: 1317-1319, 1982.

Carroll and Gheradini, "Membrane protein variations associated with *in vitro* passage of *Borrelia burgdorferi*," *Infect. Immun.*, 64:392-398, 1996.

Carter *et al.*, "A family of surface-exposed proteins of 20 kilodaltons in the genus *Borrelia*." *Infect. Immun.*, 62:2792-2799, 1994.

Casjens, Delange III, Ley, Rosa, Huang, "Linear chromosomes of Lyme disease agent spirochetes: genetic diversity and conservation of gene order," *J. Bacteriol.*, 177:2769-2780, 1995.

Demolis, Mallet, Bussereau, Jacquet, "Improved strategy for large-scale DNA sequencing using DNase I cleavage for generating radom subclones," *Biotechniques*, 18:453-457, 1995.

Donelson, "Mechanisms of antigenic variation in *Borrelia* hermsii and African trypanosomes," *J. Biol. Chem.*, 270:7783-7786, 1995.

Fuchs, Jauris, Lottspeich, Preacmursic, Wilskic, Soutschek, "Molecular analysis and expression of a *Borrelia burgdorferi* gene encoding a 22 kDa protein (pC) in *E. coli*," *Mol. Microbiol.*, 6:503-509, 1992.

Haas and Meyer, "The repertoire of silent pilus genes in Neisseria gonorrhoeae: evidence for gene conversion," *Cell*, 44:107-115, 1986.

Hagblom, Segal, Billyard, So, "Intragenic recombination leads to pilus antigenic variation in Neisseria gonorrhoeae," *Nature*, 315:156-158, 1985.

Hinnebusch, Bergstrom, Barbour, "Cloning and sequence analysis of linear plasmid telomeres of the bacterium *Borrelia burgdorferi*," *Mol. Microbiol.*, 4:811-820, 1990.

Hughes and Johnson, "Methylated DNA in *Borrelia* species," *J. Bacteriol.*, 172:6602-6604, 1990.

Johnson *et al.*, "Infection of Syrian hamsters with Lyme disease spirochetes," *J. Clin. Microbiol.*, 20:1099-1101, 1984.

Jonsson, Ilver, Falk, Pepose, Normark, "Sequence changes in the pilus subunit lead to tropism variation of Neisseria gonorrhoeae to human tissue," *Mol. Microbiol.*, 13:403-416, 1994.

Kitten and Barbour, "Juxtaposition of expressed variable antigen genes with a conserved telomere in the bacterium *Borrelia* hermsii," *Proc. Natl. Acad. Sci. USA*, 87:6077-6081, 1990.

Kitten and Barbour, "The relapsing fever agent *Borrelia* hermsii has multiple copies of its chromosome and linear plasmids," *Genetics*, 132:311-324, 1992.

Koomey, Gotschlich, Robbins, Bergstrom, Swanson, "Effects of *recA* mutations on pilus antigenic variation and phase transitions in Neisseria gonorrhoeae," *Genetics*, 117:391-398, 1987.

Kupsch, Knepper, Kuroki, Heuer, Meyer, "Variable opacity (Opa) outer membrane proteins account for the cell tropisms displayed by Neisseria gonorrhoeae for human leukocytes and epithelial cells," *EMBO. J.*, 12:641-650, 1993.

Lambden, Robertson, Watt, "Biological properties of two distinct pilus types produced by isogenic variants of Neisseria gonorrhoeae P9," *J. Bacteriol.*, 141:393-396, 1980.

Livey, Gibbs, Schuster, Dorner, "Evidence for lateral transfer and recombination in OspC variation in Lyme disease *Borrelia*," *Mol. Microbiol.*, 18:257-269, 1995.

Marconi, Konkel, Garon, "Variability of *osp* genes and gene products among species of Lyme disease spirochetes," *Infect. Immun.*, 61:2611-2617, 1993.

Marconi, Samuels, Landry, Garon, "Analysis of the distribution and molecular heterogeneity of the *ospD* gene among the Lyme disease spriochetes: evidence for lateral gene exchange," *J. Bacteriol.*, 176:4572-4582, 1994.

Margolis *et al.*, "Homology between *Borrelia burgdorferi* OspC and members of the family of *Borrelia hermsii* variable major proteins," *Gene*, 143:105-110, 1994.

Meier, Simon, Barbour, "Antigenic variation is associated with DNA rearrangements in a relapsing fever *Borrelia*," *Cell*, 41:403-409, 1985.

Meyer, Mlawer, So, "Pilus expression in Neisseria gonorrhoeae involves chromosomal rearrangement," *Cell*, 30:45-52, 1982.

Moody *et al.*, "Lyme borreliosis in laboratory animals: effect of host species and *in vitro* passage of *Borrelia burgdorferi*," *Am. J. Trop. Med. Hyg.*, 43:87-92, 1990.

Nassif, Lowry, Stenberg, O'Gaora, Ganji, So, "Antigenic variation of pilin regulates adhesion of Neisseria meningitidis to human epithelial cells," *Mol. Microbiol.*, 8:719-725, 1993.

Norris, Carter, Howell, Barbour, "Low-passage-associated proteins of *Borrelia* burgdoreferi B31: characterization and molecular cloning of OspD, a surface-exposed, plasmid-encoded lipoprotein," *Infect. Immun.*, 60:4662-4672, 1992.

Norris *et al.*, "High- and low-infectivity phenotypes of clonal populations of *in vitro*-cultured *Borrelia burgdorferi*," *Infect. Immun.*, 63:2206-2212, 1995.

Norris *et al.*, "Low-passage-associated proteins of *Borrelia burgdorferi* B31: characterization and molecular cloning of OspD, a surface exposed, plasmid-encoded lipoprotein," *Infect. Immun.*, 60:4662-4672, 1992.

Persing, Mathiesen, Podzorski, Barthold, "Genetic stability of *Borrelia burgdorferi* recovered from chronically infected immunocompetent mice," *Infect. Immun.*, 62:3521-3527, 1994.

Plasterk *et al.*, "Transposition of structural genes to an expression sequence on a linear plasmid causes antigenic variation in the bacterium *Borrelia hermsii*," *Nature*, 318:257-263, 1985.

Restrepo and Barbour, "Antigen diversity in the bacterium *B. hermsii* through 'somatic' mutations in rearranged *vmp* genes," *Cell*, 78:867-876, 1994.

Restrepo, Carter, Barbour, "Activation of a *vmp* pseudogene in *Borrelia* hermsii: an alternate mechanism of antigenic variation during relapsing fever," *Mol. Microbiol.*, 13:287-299, 1994.

Restrepo, Kitten, Carter, Infante, Barbour, "Subtelomeric expression regions of *Borrelia* hermsii linear plasmids are highly polymorphic," *Mol. Microbiol.*, 6:3299-3311, 1992.

Robertson and Meyer, "Genetic variation in pathogenic bacteria." *Trends Genet.*, 8:422-427, 1992.

Rosa, Samuels, Hogan, Stevenson, Casjens, Tilly, "Directed insertion of a selectable marker into a circular plasmid of *Borrelia burgdorferi*," *J. Bacteriol.*, 178:5946-5953, 1996.

Rosa, Schwan, Hogen, "Recombination between genes encoding major surface proteins A and B of *Borrelia burgdorferi*," *Mol. Microbiol.*, 6:3031-3040, 1992.

Rudel, Van Putten, Gibbs, Haas, Meyer, "Interaction of two variable proteins (PilE and PilC) required for pilus-mediated adherence of Neisseria gonorrhoeae to human epithelial cells," *Mol. Microbiol.*, 6:3439-3450, 1992.

Sadziene, Rosa, Thompson, Hogan, Barbour, "Antibody-resistant mutations of *Borrelia burgdorferi*: *in vitro* selection and characterization," *J. Exp. Med.*, 176:799-809, 1992.

Sambrook, Fritsch, Maniatis, "Molecular cloning: a laboratory manual," (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press), 1989.

Samuels, Mach, Garon, "Genetic transformation of the Lyme disease agent *Borrelia burgdorferi* with coumarin-resistant *gyrB*," *J. Bacteriol.*, 176:6045-6049, 1994.

Schutzer, "Lyme disease: Molecular and immunologic approaches. *In: Current communications in cell and molecular biology*," J. Inglis and J.A. Witkowski, eds. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press).

Schwann *et al.*, "Changes in infectivity and plasmid profile of the Lyme disease spirochete, *Borrelia burgdorferi*, as a result of *in vitro* cultivation." *Infect. Immun.*, 56:1831-1836. 1988a.

Schwann, Burgdorfer, Schrumpg, Karstens, "The urinary bladder, a consistent source of *Borrelia burgdorferi* in experimentally infected white-footed mice (Peromyscus leucopcus)," *J. Clin. Microbiol.*, 26:893-895, 1988b.

Schwann, Karstens, Schrumpf, Simpson, "Changes in antigenic reactivity of *Borrelia burgdorferi*, the Lyme disease spirochete, during persistent infection of mice," *Can. J. Microbiol.*, 37:450-454, 1991.

Seal, Jackson, Daniels, "Isolation of a Pseudomonas solanacearum-specific DNA probe by subtraction hybridization and construction of species-specific oligonucleotide primers for sensitive detection by the polymerase chain reaction," *Appl. Environ. Microbiol.*, 58:3751-3758, 1992.

Segal, Hagblom, Seifert, So, "Antigenic variation of gonococcal pilus involves assembly of separated silent gene segments," *Proc. Natl. Acad. Sci. USA*, 83:2177-2181, 1986.

Seifert and So, "Genetic mechanisms of bacterial antigenic variation," *Microbiol. Rev.*, 52:327-336, 1988.

Steere, "Lyme disease," *N. Engl. J. Med.*, 321:586-596, 1989.

Stevenson, Bockenstedt, Barthold, "Expression and gene sequence of outer surface protein C of *Borrelia burgdorferi* reisolated from chronically infected mice," *Infect. Immun.*, 62:3568-3571, 1994.

Stoenner, Dodd, Larsen, "Antigenic variation of *Borrelia* hermsii," *J. Exp. Med.*, 156:1297-1311, 1982.

Swanson and Koomey, "Mechanisms for variation of pili and outer membrane protein II in Neisseria gonorrhoeae," D.E. Berg and M.M. Howe, Eds. (Washington, D.C.: American Society for Microbiology).

Thiessen *et al.*, "Evolution of the *Borrelia burgdorferi* outer surface protein OspC," *J. Bacteriol.*, 177:3036-3044, 1995.

Wainwright, Pritchard, Seifert, "A conserved DNA sequence is required for efficient gonococcal pilin antigenic variation," *Mol. Microbiol.*, 13:75-87, 1994.

Walker, Howell, You, Hoffmaster, Heath, Weinstock, Norris, "Physical map of the genome of Treponema pallidum subsp. Pallidum (Nichols)." *J. Bacteriol.*, 177:1797-1804, 1995.

Wilske, Barbour, Bergstrom, Burman, Restrepo, Rosa, Schwan, Soutschek, Wallich, "Antigenic variation and strain heterogeneity in *Borrelia* spp," *Res. Microbiol.*, 143:583-596, 1992.

Wu and Tokunaga, "Biogenesis of lipoproteins in bacteria," *Curr. Top. Microbiol. Immunol.*, 125:127-157, 1986.

Xu, Kodner, Coleman, Johnson, "Correlation of plasmids with infectivity of *Borrelia burgdorferi* senso stricto type strain B31," *Infect. Immun.*, 64:3870-3876, 1996.

Xu and Johnson, "Analysis and comparison of plasmid profile of *Borrelia burgdorferi* sensu lato strains," *J. Clin. Microbiol.*, 33:2679-2685, 1995.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: Board of Regents, The University of Texas System
(B) STREET: 201 West 7th Street
(C) CITY: Austin
(D) STATE: Texas
(E) COUNTRY: USA
(F) POSTAL CODE (ZIP): 78701
(G) TELEPHONE: (512) 418-3000
(H) TELEFAX: (713) 789-2679

(ii) TITLE OF INVENTION: VMP-Like Sequences of Pathogenic Borrelia

(iii) NUMBER OF SEQUENCES: 8

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1227 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION:75..1142

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```
ACCTACACTT GTTAAAACTC TCTTTTTGAG TTAAGATGAT AACTTATACT TTTCATTATA      60

AGGAGACGAT GAAT ATG AAA AAA ATT TCA AGT GCA AGT TTA TTA ACA ACT      110
                Met Lys Lys Ile Ser Ser Ala Ser Leu Leu Thr Thr
                  1               5                  10

TTC TTT GTT TTT ATT AAT TGT AAA AGC CAA GTT GCT GAT AAG GAC GAC      158
Phe Phe Val Phe Ile Asn Cys Lys Ser Gln Val Ala Asp Lys Asp Asp
        15                  20                  25

CCA ACA AAC AAA TTT TAC CAA TCT GTC ATA CAA TTA GGT AAC GGA TTT      206
Pro Thr Asn Lys Phe Tyr Gln Ser Val Ile Gln Leu Gly Asn Gly Phe
    30                  35                  40
```

```
CTT GAT GTA TTC ACA TCT TTT GGT GGG TTA GTA GCA GAG GCT TTT GGA      254
Leu Asp Val Phe Thr Ser Phe Gly Gly Leu Val Ala Glu Ala Phe Gly
 45                  50                  55                  60

TTT AAA TCA GAT CCA AAA AAA TCT GAT GTA AAA ACC TAT TTT ACT ACT      302
Phe Lys Ser Asp Pro Lys Lys Ser Asp Val Lys Thr Tyr Phe Thr Thr
                 65                  70                  75

GTA GCT GCC AAA TTG GAA AAA ACA AAA ACC GAT CTT AAT AGT TTG CCT      350
Val Ala Ala Lys Leu Glu Lys Thr Lys Thr Asp Leu Asn Ser Leu Pro
             80                  85                  90

AAG GAA AAA AGC GAT ATA AGT AGT ACG ACG GGG AAA CCA GAT AGT ACA      398
Lys Glu Lys Ser Asp Ile Ser Ser Thr Thr Gly Lys Pro Asp Ser Thr
         95                 100                 105

GGT TCT GTT GGA ACT GCC GTT GAG GGG GCT ATT AAG GAA GTT AGC GAG      446
Gly Ser Val Gly Thr Ala Val Glu Gly Ala Ile Lys Glu Val Ser Glu
    110                 115                 120

TTG TTG GAT AAG CTG GTA AAA GCT GTA AAG ACA GCT GAG GGG GCT TCA      494
Leu Leu Asp Lys Leu Val Lys Ala Val Lys Thr Ala Glu Gly Ala Ser
125                 130                 135                 140

AGT GGT ACT GCT GCA ATT GGA GAA GTT GTG GCT GAT GCT GAT GCT GCA      542
Ser Gly Thr Ala Ala Ile Gly Glu Val Val Ala Asp Ala Asp Ala Ala
                145                 150                 155

AAG GTT GCT GAT AAG GCG AGT GTG AAG GGG ATT GCT AAG GGG ATA AAG      590
Lys Val Ala Asp Lys Ala Ser Val Lys Gly Ile Ala Lys Gly Ile Lys
            160                 165                 170

GAG ATT GTT GAA GCT GCT GGG GGG AGT GAA AAG CTG AAA GCT GTT GCT      638
Glu Ile Val Glu Ala Ala Gly Gly Ser Glu Lys Leu Lys Ala Val Ala
        175                 180                 185

GCT GCT AAA GGG GAG AAT AAT AAA GGG GCA GGG AAG TTG TTT GGG AAG      686
Ala Ala Lys Gly Glu Asn Asn Lys Gly Ala Gly Lys Leu Phe Gly Lys
    190                 195                 200

GCT GGT GCT GCT GCT CAT GGG GAC AGT GAG GCT GCT AGC AAG GCG GCT      734
Ala Gly Ala Ala Ala His Gly Asp Ser Glu Ala Ala Ser Lys Ala Ala
205                 210                 215                 220

GGT GCT GTT AGT GCT GTT AGT GGG GAG CAG ATA TTA AGT GCG ATT GTT      782
Gly Ala Val Ser Ala Val Ser Gly Glu Gln Ile Leu Ser Ala Ile Val
                225                 230                 235

ACG GCT GCT GAT GCG GCT GAG CAG GAT GGA AAG AAG CCT GAG GAG GCT      830
Thr Ala Ala Asp Ala Ala Glu Gln Asp Gly Lys Lys Pro Glu Glu Ala
            240                 245                 250

AAA AAT CCG ATT GCT GCT GCT ATT GGG GAT AAA GAT GGG GGT GCG GAG      878
Lys Asn Pro Ile Ala Ala Ala Ile Gly Asp Lys Asp Gly Gly Ala Glu
        255                 260                 265
```

```
TTT GGT CAG GAT GAG ATG AAG AAG GAT GAT CAG ATT GCT GCT GCT ATT        926
Phe Gly Gln Asp Glu Met Lys Lys Asp Asp Gln Ile Ala Ala Ala Ile
    270                 275                 280

GCT TTG AGG GGG ATG GCT AAG GAT GGA AAG TTT GCT GTG AAG GAT GGT        974
Ala Leu Arg Gly Met Ala Lys Asp Gly Lys Phe Ala Val Lys Asp Gly
285                 290                 295                 300

GAG AAA GAG AAG GCT GAG GGG GCT ATT AAG GGA GCT GCT GAG TCT GCA       1022
Glu Lys Glu Lys Ala Glu Gly Ala Ile Lys Gly Ala Ala Glu Ser Ala
                305                 310                 315

GTT CGC AAA GTT TTA GGG GCT ATT ACT GGG CTA ATA GGA GAC GCC GTG       1070
Val Arg Lys Val Leu Gly Ala Ile Thr Gly Leu Ile Gly Asp Ala Val
            320                 325                 330

AGT TCC GGG CTA AGG AAA GTC GGT GAT TCA GTG AAG GCT GCT AGT AAA       1118
Ser Ser Gly Leu Arg Lys Val Gly Asp Ser Val Lys Ala Ala Ser Lys
        335                 340                 345

GAA ACA CCT CCT GCC TTG AAT AAG TGATTTAATT AAGTGTATGG ACACGACTAT      1172
Glu Thr Pro Pro Ala Leu Asn Lys
    350                 355

GCCCTCATGA TTGAGGAAAT AGTCGAGAGA TATATATACT AAAAGATAAT AAATA          1227
```

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 356 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```
Met Lys Lys Ile Ser Ser Ala Ser Leu Leu Thr Thr Phe Phe Val Phe
 1               5                  10                  15

Ile Asn Cys Lys Ser Gln Val Ala Asp Lys Asp Asp Pro Thr Asn Lys
            20                  25                  30

Phe Tyr Gln Ser Val Ile Gln Leu Gly Asn Gly Phe Leu Asp Val Phe
        35                  40                  45

Thr Ser Phe Gly Gly Leu Val Ala Glu Ala Phe Gly Phe Lys Ser Asp
    50                  55                  60

Pro Lys Lys Ser Asp Val Lys Thr Tyr Phe Thr Thr Val Ala Ala Lys
65                  70                  75                  80

Leu Glu Lys Thr Lys Thr Asp Leu Asn Ser Leu Pro Lys Glu Lys Ser
                85                  90                  95
```

Asp Ile Ser Ser Thr Thr Gly Lys Pro Asp Ser Thr Gly Ser Val Gly
100 105 110

Thr Ala Val Glu Gly Ala Ile Lys Glu Val Ser Glu Leu Leu Asp Lys
115 120 125

Leu Val Lys Ala Val Lys Thr Ala Glu Gly Ala Ser Ser Gly Thr Ala
130 135 140

Ala Ile Gly Glu Val Val Ala Asp Ala Asp Ala Ala Lys Val Ala Asp
145 150 155 160

Lys Ala Ser Val Lys Gly Ile Ala Lys Gly Ile Lys Glu Ile Val Glu
165 170 175

Ala Ala Gly Gly Ser Glu Lys Leu Lys Ala Val Ala Ala Ala Lys Gly
180 185 190

Glu Asn Asn Lys Gly Ala Gly Lys Leu Phe Gly Lys Ala Gly Ala Ala
195 200 205

Ala His Gly Asp Ser Glu Ala Ala Ser Lys Ala Ala Gly Ala Val Ser
210 215 220

Ala Val Ser Gly Glu Gln Ile Leu Ser Ala Ile Val Thr Ala Ala Asp
225 230 235 240

Ala Ala Glu Gln Asp Gly Lys Lys Pro Glu Glu Ala Lys Asn Pro Ile
245 250 255

Ala Ala Ala Ile Gly Asp Lys Asp Gly Gly Ala Glu Phe Gly Gln Asp
260 265 270

Glu Met Lys Lys Asp Asp Gln Ile Ala Ala Ala Ile Ala Leu Arg Gly
275 280 285

Met Ala Lys Asp Gly Lys Phe Ala Val Lys Asp Gly Glu Lys Glu Lys
290 295 300

Ala Glu Gly Ala Ile Lys Gly Ala Ala Glu Ser Ala Val Arg Lys Val
305 310 315 320

Leu Gly Ala Ile Thr Gly Leu Ile Gly Asp Ala Val Ser Ser Gly Leu
325 330 335

Arg Lys Val Gly Asp Ser Val Lys Ala Ala Ser Lys Glu Thr Pro Pro
340 345 350

Ala Leu Asn Lys
355

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7766 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:
(A) NAME/KEY: modified_base
(B) LOCATION:127
(D) OTHER INFORMATION:/note= "R = A or G"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GCTGGTGCTG GTAATGCTGG GGACAGTGAG GCTGCTAGCA AGGCGGCTGG TGCTGTTAGT 60

GCTGTTAGTG GGGAGCAGAT ATTAAGTGCG ATTGTTAAGG CTGCTGGTGA GGCTGCGCAG 120

GATGGARAGA AGCCTGGGGA GGCTAAAAAT CCGATTGCTG CTGCTATTGG GAAGGGTAAT 180

GAGGATGGTG CGGAGTTTAA GGATGAGATG AAGAAGGATG ATCAGATTGC TGCTGCTATT 240

GCTTTGAGGG GGATGGCTAA GGATGGAAAG TTTGCTGTGA AGAATGATGA GAAAGGGAAG 300

GCTGAGGGGG CTATTAAGGG AGCTGGCGAG TTGTTGGATA AGCTGGTAAA AGCTGTAAAG 360

ACAGCTGAGG GGGCTTCAAG TGGTACTGCT GCAATTGGAG AAGTTGTGGC TGATGATAAT 420

GCTGCGAAGG TTGCTGATAA GGCGAGTGTG AAGGGGATTG CTAAGGGGAT AAAGGAGATT 480

GTTGAAGCTG CTGGGGGGAG TAAAAAGCTG AAAGTTGCTG CTGCTAAAGA GGGCAATGAA 540

AAGGCAGGGA AGTTGTTTGG GAAAGTTGAT GCTGCTCATG CTGGGGACAG TGAGGCTGCT 600

AGCAAGGCGG CTGGTGCTGT TAGTGCTGTT AGTGGGGAGC AGATATTAAG TGCGATTGTT 660

AAGGCTGCTG GTGCGGCTGC TGGTGATCAG GAGGGAAAGA AGCCTGGGGA TGCTAAAAAT 720

CCGATTGCTG CTGCTATTGG GAAGGGTGAT GCGGAGAATG GTGCGGAGTT TAATCATGAT 780

GGGATGAAGA AGGATGATCA GATTGCTGCT GCTATTGCTT TGAGGGGGAT GGCTAAGGAT 840

GGAAAGTTTG CTGTGAAGAG TGGTGGTGGT GAGAAAGGGA AGGCTGAGGG GGCTATTAAG 900

GGAGCTGCTG AGTTGTTGGA TAAGCTGGTA AAAGCTGTAA AGACAGCTGA GGGGGCTTCA 960

AGTGGTACTG ATGCAATTGG AGAAGTTGTG GCTAATGCTG GTGCTGCAAA GGTTGCTGAT 1020

AAGGCGAGTG TGACGGGGAT TGCTAAGGGG ATAAAGGAGA TTGTTGAAGC TGCTGGGGGG 1080

AGTGAAAAGC TGAAAGTTGC TGCTGCTACA GGGGAGAGTA ATAAAGGGGC AGGGAAGTTG 1140

TTTGGGAAGG CTGGTGCTGG TGCTAATGCT GGGGACAGTG AGGCTGCTAG CAAGGCGGCT 1200

GGTGCTGTTA GTGCTGTTAG TGGGGAGCAG ATATTAAGTG CGATTGTTAA GGCTGCTGAT 1260

GCGGCTGATC AGGAGGGAAA GAAGCCTGGG GATGCTACAA ATCCGATTGC TGCTGCTATT 1320

GGGAAGGGTA ATGAGGAGAA TGGTGCGGAG TTTAAGGATG AGATGAAGAA GGATGATCAG 1380

ATTGCTGCTG CTATTGCTTT GAGGGGGATG GCTAAGGATG GAAAGTTTGC TGTGAAGGAT 1440

GGTGGTGAGA AAGGGAAGGC TGAGGGGGCT ATTAAGGGAG CTGCTGAGTT GTTGGATAAG 1500

CTGGTAAAAG CTGTAAAGAC AGCTGAGGGG GCTTCAAGTG GTACTGATGC AATTGGAGAA 1560

GTTGTGGATA ATGCTGCGAA GGCTGCTGAT AAGGCGAGTG TGACGGGGAT TGCTAAGGGG 1620

ATAAAGGAGA TTGTTGAAGC TGCTGGGGGG AGTGAAAAGC TGAAAGTTGC TGCTGCTACA 1680

GGGGAGAATA ATAAAGAGGC AGGGAAGTTG TTTGGGAAGG CTGGTGCTGA TGCTAATGGG 1740

GACAGTGAGG CTGCTAGCAA GGCGGCTGGT GCTGTTAGTG CTGTTAGTGG GGAGCAGATA 1800

TTAAGTGCGA TTGTTAAGGC TGCGGCTGCT GGTGCGGCTG ATCAGGATGG AGAGAAGCCT 1860

GGGGATGCTA AAAATCCGAT TGCTGCTGCT ATTGGGAAGG GTAATGCGGA TGATGGTGCG 1920

GATTTTGGTG ATGGGATGAA GAAGGATGAT CAGATTGCTG CTGCTATTGC TTTGAGGGGG 1980

ATGGCTAAGG ATGGAAAGTT TGCTGTGAAG AAGGATGAGA AAGGGAAGGC TGAGGGGCCT 2040

ATTAAGGGAG CTAGCGAGTT GTTGGATAAG CTGGTAAAAG CTGTAAAGAC AGCTGAGGGG 2100

GCTTCAAGTG GTACTGCTGC AATTGGAGAA GTTGTGGATA ATGCTGCGAA GGCTGCTGAT 2160

AAGGATAGTG TGACGGGGAT TGCTAAGGGG ATAAAGGAGA TTGTTGAAGC TGCAGGGGGG 2220

AGTGAAAAGC TGAAAGTTGC TGCTGCTAAA GGGGAGAATA ATAAAGGGGC AGGGAAGTTG 2280

TTTGGGAAGG CTGGTGCTAA TGCTCATGGG GACAGTGAGG CTGCTAGCAA GGCGGCTGGT 2340

GCTGTTAGTG CTGTTAGTGG GGAACAGATA TTAAGTGCGA TTGTTAAGGC TGCTGGTGAG 2400

GCTGCTGGTG ATCAGGAGGG AAAGAAGCCT GAGGAGGCTA AAAATCCGAT TGCTGCTGCT 2460

ATTGGGGATA AAGATGGGGA TGCGGAGTTT AATCAGGATG GGATGAAGAA GGATGATCAG 2520

ATTGCTGCTG CTATTGCTTT GAGGGGGATG GCTAAGGATG GAAAGTTTGC TGTGAAGGAT 2580

GGTGGTGAGA AAGAGAAGGC TGAGGGGGCT ATTAAAGGAG TTAGCGAGTT GTTGGATAAG 2640

CTGGTAAAAG CTGTAAAGAC AGCTGAGGGG GCTTCAAGTG GTACTGCTGC AATTGGAGAA 2700

GTTGTGGCTG ATGCTGCTAA GGTTGCTGAT AAGGCGAGTG TGACGGGGAT TGCTAAGGGG 2760

ATAAAGGAGA TTGTTGAAGC TGCTGGGGAC AGTGAGGCTG CTAGCAAGGC AGCTGGTGCT 2820

GTTAGTGCTG TTAGTGGGGA GCAGATATTA AGTGCGATTG TTAAGGCTGC GGCTGCTGGT 2880

GCGGCTGAGC AGGATGGAGA GAAGCCTGCA GAGGCTAAAA ATCCGATTGC TGCTGCTATT 2940

GGGAAGGGTG ATGGGGATGC GGATTTTGGT GAGGATGGGA TGAAGAAGGA TGATCAGATT 3000

GCTGCTGCTA TTGCTTTGAG GGGGATGGCT AAGGATGGAA AGTTTGCTGT GAAGAATGAT 3060

GAGAAAGGGA AGGCTGAGGG GGCTATTAAG GGAGCTGCTG CAATTGGAGA AGTTGTGGAT 3120

AATGCTGGTG CTGCGAAGGC TGCTGATAAG GATAGTGTGA AGGGGATTGC TAAGGGGATA 3180

AAGGAGATTG TTGAAGCTGC TGGGGGGAGT GAAAAGCTGA AAGCTGCTGC TGCTGAAGGG 3240

GAGAATAATA AAAAGGCAGG GAAGTTGTTT GGGAAAGTTG ATGGTGCTGC TGGGGACAGT 3300

GAGGCTGCTA GCAAGGCGGC TGGTGCTGTT AGTGCTGTTA GTGGGGAGCA GATATTAAGT 3360

GCGATTGTTA AGGCTGCTGG TGAGGCTGAG CAGGATGGAG AGAAGCCTGA GGATGCTAAA 3420

AATCCGATTG CTGCTGCTAT TGGGAAGGGT AATGGGGATG GTGCGGAGTT TGATCAGGAT 3480

GAGATGAAGA AGGATGATCA GATTGCTGCT GCTATTGCTT TGAGGGGGAT GGCTAAGGAT 3540

GGAAAGTTTG CTGTGAAGGG TAATAATGAG AAAGAGAAGG CTGAGGGGGC TATTAAAGAA 3600

GTTAGCGAGT TGTTGGATAA GCTGGTAACA GCTGTAAAGA CAGCTGAGGG GGCTTCAAGT 3660

GGTACTGATG CAATTGGAGA AGTTGTGGAT AATGATGCTA AGGTTGCTGA TAAGGCGAGT 3720

GTGACGGGGA TTGCTAAGGG GATAAAGGAG ATTGTTGAAG CTGCTAGGGG GAGTGAAAAG 3780

CTGAAAGTTG CTGCTGCTAA AGAGGGCAAT GAAAAGGCAG GGAAGTTGTT TGGGAAGGCT 3840

GGTGCTAATG CTCATGGGGA CAGTGAGGCT GCTAGCAAGG CGGCTGGTGC TGTTAGTGCT 3900

GTTAGTGGGG AGCAGATATT AAGTGCGATT GTTAAGGCTG CGGATGCGGC TGAGCAGGAT 3960

GGAAAGAAGC CTGCAGATGC TACAAATCCG ATTGCTGCTG CTATTGGGAA TAAAGATGAG 4020

GATGCGGATT TTGGTGATGG GATGAAGAAG GATGATCAGA TTGCTGCTGC TATTGCTTTG 4080

AGGGGGATGG CTAAGGATGG AAAGTTTGCT GTGAAGGGTA ATAATGAGAA AGGGAAGGCT 4140

GAGGGGGCTT CAAGTGGTAC TGATGCAATT GGAGAAGTTG TGGATAATGA TGCGAAGGCT 4200

GCTGATAAGG CGAGTGTGAC GGGGATTGCT AAGGGGATAA AGGAGATTGT TGAAGCTGCT 4260

GGGGGGAGTG AAAAGCTGAA AGCTGTTGCT GCTGCTACAA GGGAGAATAA TAAAGAGGCA 4320

GGGAAGTTGT TTGGGAAAGT TGATGATGCT CATGCTGGGG ACAGTGAGGC TGCTAGCAAG 4380

GCGGCTGGTG CTGTTAGTGC TGTTAGTGGG GAGCAGATAT TAAGTGCGAT TGTTACGGCT 4440

GCGGCTGCTG GTGAGCAGGA TGGAGAGAAG CCTGCAGAGG CTACAAATCC GATTGCTGCT 4500

GCTATTGGGA AGGGTAATGA GGATGGTGCG GATTTTGGTA AGGATGAGAT GAAGAAGGAT 4560

GATCAGATTG CTGCTGCTAT TGCTTTGAGG GGGATGGCTA AGGATGGAAA GTTTGCTGTG 4620

AAGAGTAATG ATGGTGAGAA AGGGAAGGCT GAGGGGGCTA TTAAGGAAGT TAGCGAGTTG 4680

TTGGATAAGC TGGTAAAAGC TGTAAAGACA GCTGAGGGGG CTTCAAGCGG TACTGATGCA 4740

ATTGGAGAAG TTGTGGCTAA TGCTGGTGCT GCGAAGGCTG CTGATAAGGC GAGTGTGACG 4800

GGGATTGCTA AGGGGATAAA GGAGATTGTT GAAGCTGCTG GGGGGAGTAA AAAGCTGAAA 4860

GCTGCTGCTG CTGAAGGGGA GAATAATAAA AAGGCAGGGA AGTTGTTTGG GAAGGCTGGT 4920

GCTGGTGCTG GTGCTAATGG GGACAGTGAG GCTGCTAGCA AGGCGGCTGG TGCTGTTAGT 4980

GCTGGTTAGT GTGGGGAGCA GATATTAAGT GCGATTGTTA CGGCTGCTGG TGCGGCTGCT 5040

AGTGAGGCTG ATCAGGAGGG AAAGAAGCCT GCAGATGCTA CAAATCCGAT TGCTGCTGCT 5100

ATTGGGAAGG GTGATGCGGA GAATGGTGCG GATTTTGGTG ATGGGATGAA GAAGGATGAT 5160

CAGATTGCTG CTGCTATTGC TTTGAGGGGG ATGGCTAAGG ATGGAAAGTT TGCTGTGAAG 5220

AATGATGATG AGAAAGGGAA GGCTGAGGGG GCTATTAAGG GAGCTAGCGA GTTGTTGGAT 5280

AAGCTGGTAA CAGCTGTAAA GACAGCTGAG GGGGCTTCAA GTGGTACTGA TGCAATTGGA 5340

GAAGTTGTGG CTGATGCTGC GAAGGCTGCT GATAAGGATA GTGTGAAGGG GATTGCTAAG 5400

GGGATAAAGG AGATTGTTGA AGCTGCTGGG GGGAGTGAAA AGCTGAAAGT TGCTGCTGCT 5460

AAAGAGGGCA ATGAAAAGGC AGGGAAGTTG TTTGGGAAGG TTGGTGATGC TGCTCATGCT 5520

GGGGACAGTG AGGCTGCTAG CAAGGCGGCT GGTGCTGTTA GTGCTGTTAG TGGGGAGCAG 5580

ATATTAAGTG CGATTGTTAC GGCTGCTGGT GCGGCTGAGC AGGAGGGAAA GAAGCCTGCA 5640

GAGGCTAAAA ATCCGATTGC TGCTGCTATT GGGAAGGGTA ATGAGAATGG TGCGGAGTTT 5700

AAGGATGAGA TGAAGAAGGA TGATCAGATT GCTGCTGCTA TTGCTTTGAG GGGGATGGCT 5760

AAGGATGGAA AGTTTGCTGT GAAGAAGGAT AATAATGAGA AAGGGAAGGC TGAGGGGGCT 5820

ATTAAGGAAG TTAGCGAGTT GTTGGATAAG CTGGTAACAG CTGTAAAGAC AGCTGAGGAG 5880

GCTTCAAGTG GTACTGCTGC AATTGGAGAA GTTGTGGCTG ATGATGCTGC TGCGAAGGCT 5940

GCTGATAAGG AGAGTGTGAA GGGGATTGCT AAGGGGATAA AGGAGATTGT TGAAGCTGCT 6000

GGGGGGAGTA AAAAGCTGAA AGTTGCTGCT GCTACAGGGG AGAATAATAA AAAGGCAGGG 6060

AAGTTGTTTG GGAAAGTTGA TGCTGGTAAT GCTGGGGACA GTGAGGCTGC TAGCAAGGCG 6120

GCTGGTGCTG TTAGTGGGGA GCAGATATTA AGTGCGATTG TTAAGGCTGC TGGTGCGGCT 6180

GCTGGTGATC AGGAGGGAAA GAAGCCTGGG GATGCTAAAA ATCCGATTGC TGCTGCTATT 6240

GGGAAGGGTG ATGCGGAGAA TGGTGCGGAG TTTGATCATG AGATGAAGAA GGATGATCAG 6300

ATTGCTGCTG CTATTGCTTT GAGGGGGATG GCTAAGGATG GAAAGTTTGC TGTGAAGAGT 6360

GGTGATGAGA AAGGGAAGGC TGAGGGGGCT ATTAAGGGAG CTAGCGAGTT GTTGGATAAG 6420

CTGGTAAAAG CTGTGTAAAG ACAGCTGAGG GGGCTTCAAG TGGTACTGAT GCAATTGGAG 6480

AAGTTGTGGC TAATGATGCT GCTGCGAAGG TTGCTGATAA GGAGAGTGTG ACGGGGATTG 6540

CTAAGGGGAT AAAGGAGATT GTTGAAGCTG CTGGGGGGAG TGAAAAGCTG AAAGTTGCTG 6600

CTGCTACAAG GGAGAATAAT AAAAAGGCAG GGAAGTTGTT TGGGAAAGCT GGTGATGCTG 6660

CTAATGGGGA CAGTGAGGCT GCTAGCAAGG CGGCTGGTGC TGTTAGTGCT GTTAGTGGGG 6720

AGCAGATATT AAGTGCGATT GTTACGGCTG CAGCTGCTGG TGCGGCTGAG CAGGAGGGAA 6780

AGAAGCCTGA GGAGGCTAAA AATCCGATTG CTGCTGCTAT TGGGAAGGGT AATGCGGATG 6840

ATGGTGCGGA GTTTAATAAG GAGGGGATGA AGAAGGATGA TCAGATTGCT GCTGCTATTG 6900

CTTTGAGGGG GATGGCTAAG GATGGAAAGT TTGCTGTGAA GAGTGGTGGT GAGAAAGGGA 6960

AGGCTGAGGG GGCTATTAAA GAAGTTAGCG AGTTGTTGGA TAAGCTGGTA ACAACTGTAA 7020

AGACAGCTGA GGGGGCTTCA AATGGTACTG ATGCAATTGG AAAAGTTGTG GATAATAATA 7080

ATGCTGCGAA GGCTGCTGAT AAGGCGAGTG TGACGGGGAT TGCTAAGGGG ATAAAGGAGA 7140

TTGTTGAAGC TGCTGGGGGG AGTAGTGAAA AGCTGAAAGC TGTTGCTGCT GCTAAAGGGG 7200

AGAGCAATGA AAAGGCAGGG AAGTTGTTTG GGAAGGCTGG TGCTGCTGCT GGGGACAGTG 7260

AGGCTGCTAG CAAGGCGGCT GGTGCTGTTA GTGCTGTTAG TGGGGAGCAG ATATTAAGTG 7320

CGATTGTTAA GGCTGCTGGT GCGGCTGATC AGGAGGGAAA GAAGCCTGAG GATGCTAAAA 7380

ATCCGATTGC TGCTGCTATT GGGGATAAAG ATGGGGGTGC GGAGTTTAAT CATGAGATGA 7440

AGAAGGATGA TCAGATTGCT GCTGCTATTG CTTTGAGGGG GATGGCTAAG GATGGAAAGT 7500

TTGCTGTGAA GAGTGGTGGT GGTGAGAAAG AGAAGGCTGA GGGGGCTATT AAAGAAGTTA 7560

GCGAGTTGTT GGATAAGCTG GTAAAAGCTG TGTAAAGACA GCTGAGGGGG CTTCAAGTGG 7620

TACTGATGCA ATTGGAGAAG TTGTGGCTGA TAATAGTGCT GCGAAGGCTG CTGATGAGGC 7680

GAGTGTGACG GGGATTGCTA AGGGAATAAA GGAGATTGTT GAAGCTGCTG GGGGGAGTGA 7740

AAAGCTGAAA GTTGCTGCTG CTGCAG 7766

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

AGTACGGGGA AACCAG 16

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

CTTTGCGAAC GCAGACTCAG CA 22

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

AGTGGGGAGA TATTAAGTGC G 21

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

CTTTGCGAAC GCAGACTCAG CA 22

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

GAGGGGGCTA TTAAGGA 17

CLAIMS

1. A purified nucleic acid segment encoding aVMP-like peptide of *Borrelia*.

2. The nucleic acid segment of claim 1 further defined as encoding a peptide having the amino acid sequence of SEQ ID NO:2.

3. The nucleic acid segment of claim 1 further defined as comprising the nucleic acid sequence of SEQ ID NO:1, or the complement thereof, or a sequence of at least 15 nucleotides which hybridizes to the sequence of SEQ ID NO:1 under conditions of high stringency.

4. The nucleic acid segment of claim 1 further defined as an RNA segment.

5. A DNA segment comprising an isolated *Borrelia* that encodes a Vmp-like sequence.

6. The DNA segment of claim 5 comprising an isolated *Borrelia* gene that includes a contiguous nucleic acid sequence of at least 16 nucleotides from SEQ ID NO:1.

7. The DNA segment of claim 5 comprising an isolated *Borrelia* gene that encodes a protein from about 50 to about 150 amino acids in length.

8. The DNA segment of claim 5 comprising an isolated *Borrelia* gene that encodes a protein from about 100 to about 200 amino acids in length.

9. The DNA segment of claim 5 comprising an isolated *Borrelia* gene that encodes a protein from about 200 to about 400 amino acids in length.

10. A recombinant host cell comprising the DNA segment of claim 5.

11. The recombinant host cell further defined as an *E. coli* cell.

12. A method of using a DNA segment that encodes an isolated *Borrelia* Vmp-like protein gene, comprising the steps:

(a) preparing a recombinant vector in which the Vmp-like protein gene DNA segment of claim 5 is positioned under the control of a promoter;

(b) introducing said recombinant vector into a host cell;

(c) culturing said host cell under conditions effective to allow expression of the encoded protein or peptide; and

(d) collecting said expressed Vmp-like protein or peptide.

13. A method for diagnosing Lyme disease, comprising identifying a Vmp-like nucleic acid segment having a nucleotide sequence as set forth in claim 3 or a Vmp-like protein or peptide encoded by said nucleic acid segment which is present in a clinical sample from a patient suspected of having Lyme disease.

14. An isolated polypeptide which specifically binds with antibodies raised against a polypeptide having at least the amino acid sequence of SEQ ID NO:2.

15. An isolated polypeptide comprising the amino acid sequence of SEQ ID NO:2.

16. The polypeptide of claim 14 that has the amino acid sequence of SEQ ID NO:2.

17. The polypeptide of claim 14 further defined as having at least 85% homology to the amino acid sequence of SEQ ID NO:2.

18. A protein composition comprising the polypeptide of claim 14.

19. The composition of claim 18 further defined as comprising a physiologically acceptable excipient.

20. A purified antibody that specifically binds to the *Borrelia* Vmp-like protein or peptide of claim 14.

21. The antibody of claim 20 wherein the antibody is linked to a detectable label.

22. A method of generating an immune response, comprising administering to a mammal a pharmaceutical composition comprising an immunologically effective amount of the composition of claim 19.

23. A method for detecting *Borrelia* Vmp-like protein in a biological sample, comprising the steps of:

(a) obtaining a biological sample suspected of containing a *Borrelia* Vmp-like protein;

(b) contacting said sample with a first antibody that specifically binds to the protein under conditions effective to allow formation of an immune complex; and

(c) detecting the immune complex so formed.

24. An immunodetection kit comprising, in suitable container means, one or more proteins as defined by claim 14, or an antibody that binds to a protein as defined in claim 14, and an immunodetection agent.

25. A polypeptide encoded by any of *Borrelia* gene segments *vls2-vls16* wherein said segment has the nucleic acid sequence of SEQ ID NO:1 and SEQ ID NO:3 from position 712 to 1293 or from position 1294 to 1869 or from position 1870 to 2439 or from position 2440 to 3009 or from 3010 to 3483 or from position 3484 to 3990 or from 3991 to 4548 or from position 4549 to 5058 or from 5059 to 5652 or from 4653 to 6219 or from position 6220 to 6789 or from position 6846 to 7373 or from 7274 to 7946 or from 7947 to 8000 or wherein said polypeptide is encoded by a repeat gene segment from position 1293 to 1309 or from position 1869 to 1885 or to position 2439 to 2456 or position 3009 to 3025 or position 3483 to 3499 or from position 3990 to 4006 or from position 4548 to 4557 or from position 5058 to 5074 or

from position 5652 to 5668 or from position 6219 to 6253 or from position 6789 to 6805 or from position 7373 to 7389 or from position 7946 to 7962 of SEQ ID NO:1 and SEQ ID NO:3.

26. The polypeptide of claim 25 characterized as specifically binding with one or more of antibodies raised against a polypeptide encoded by SEQ ID NO:1 and SEQ ID NO:3.

27. An isolated nucleic acid segment obtained by the amplification of a *Borrelia* nucleic acid with a set of primers in accordance with SEQ ID NO: 4 and SEQ ID NO:5

28. The nucleic acid segment of claim 27 wherein the *Borrelia* DNA is *B. burgdorferi*.

29. The polypeptide of claim 25 encoded as a genetic recombination of the vls2-vls16 gene segments from SEQ ID NO:1 and SEQ ID NO:3 wherein said segments include nucleic acid segments from position 712 to 1293 or from position 1294 to 1869 or from position 1870 to 2439 or from position 2440 to 3009 or from 3010 to 3483 or from position 3484 to 3990 or from 3991 to 4548 or from position 4549 to 5058 or from 5059 to 5652 or from 4653 to 6219 or from position 6220 to 6789 or from position 6846 to 7373 or from 7274 to 7946 or from 7947 to 8000 or from repeat gene segments from nucleic acid position 1293 to 1309 or from position 1869 to 1885 or to position 2439 to 2456 or position 3009 to 3025 or position 3483 to 3499 or from position 3990 to 4006 or from position 4548 to 4557 or from position 5058 to 5074 or from position 5652 to 5668 or from position 6219 to 6253 or from position 6789 to 6805 or from position 7373 to 7389 or from position 7946 to 7962 of SEQ ID NO:1 and SEQ ID NO:3.

30. A recombinant gene segment produced by genetic recombination from any of the vls2-vls16 gene segments of claim 29.

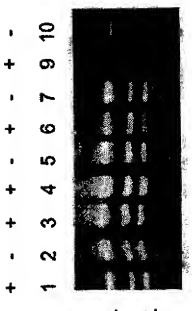

FIG.1A

FIG.1B

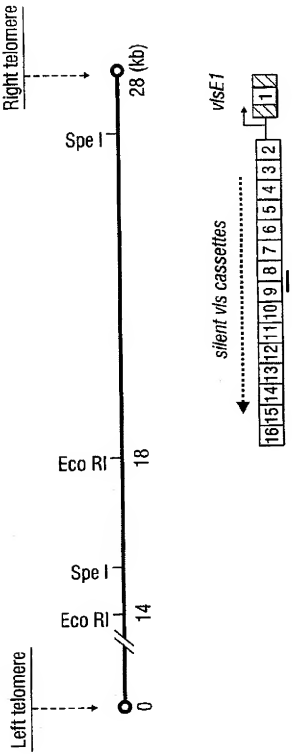
Right telomere
28 (kb)
Spe I
Multiple Pst I, Rsa I and Sau 3A1 sites
vlsE1
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
silent vls cassettes
pJRZ53
Eco RI
18
Spe I
Eco RI
14
Left telomere
0
λDASH-Bb12

FIG. 2A

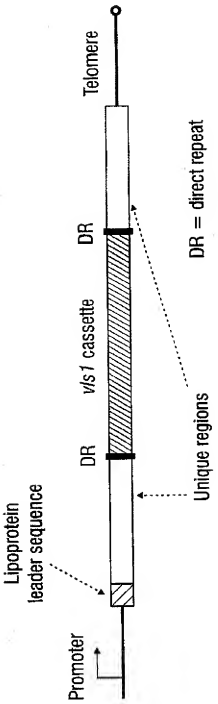
Telomere
DR
vls1 cassette
DR
Lipoprotein
leader sequence
Promoter
DR = direct repeat
Unique regions

FIG. 2B

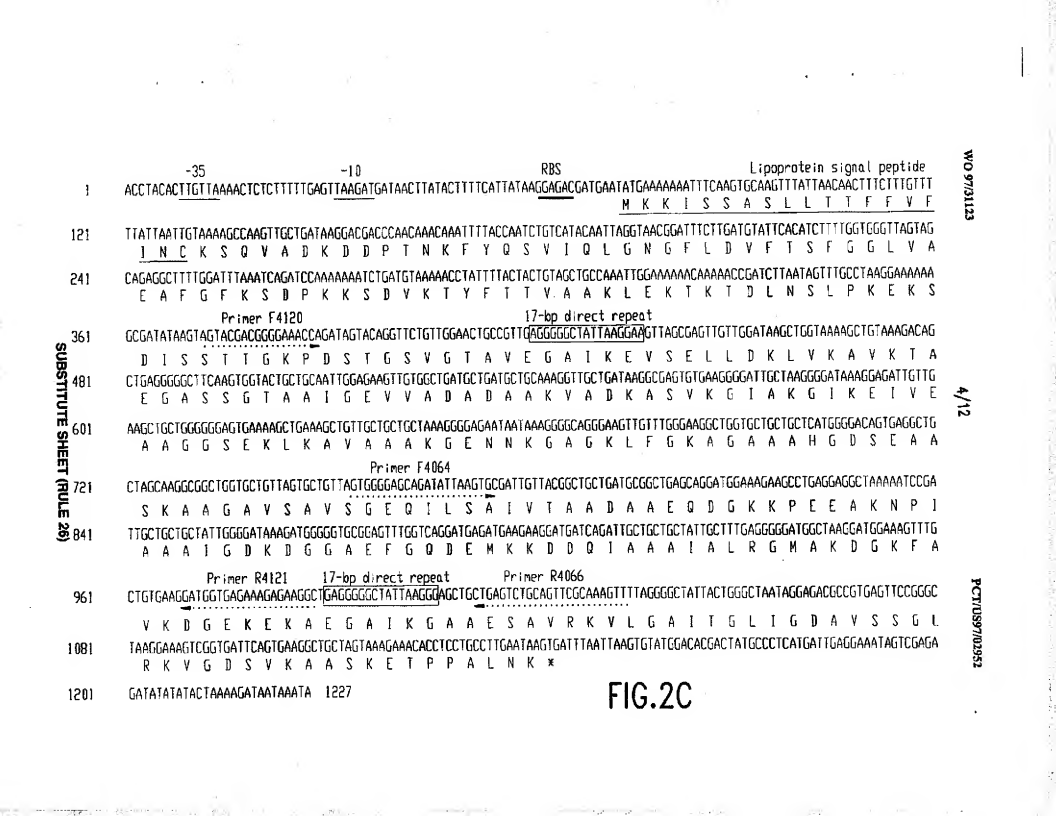
-35
-10
RBS
Lipoprotein signal peptide
Primer F4120
17-bp direct repeat
Primer F4064
Primer R4121
17-bp direct repeat
Primer R4066

FIG.2C

VisE1 1 MKKISSASLLTTFFVFINCKSGVADKDPTNKFYGSVIQLGNGFLDVFTGFGDLYAEAFGFKSDPKKSDVKTYFTTVAAKLEKTKTDLNSLPKEKSDISS
Vmp17 1 MRKRISAIIMTLFMVLVSCKSGGVA.EDPKTVYLTSTANLGKGFLDVFVTFGDWTGAFGIKADTKKSDIGKYFTDIESTMTSVKKKLQ.....D
VisE1 101 TTGKPDSTGSGVGTAVEGAIKEVSELLDKLVKAVKTAEGASSGTAAIGEVVADADAHKVADKASYKGIANGINEIVEAAGGSEKLKAVAAAKGENNKGAGK
Vmp17 90 EVAKNGNYPKVKTAVD....EFVAILGKIEKGAKEASKGATGDVIIGNTVKNGDAV.PGEATSVNSQLVKGIIKEIVGNVLKEGKADA.DATKDDSKKDIGK
VisE1 201 LFGKAGAAAHGDSEAASKAAGAVSAVSGEQILSAIV.....TAADAAEDDGKKPEEAKNPIAAAIGDKDGGAEFGGDEMKKDDDIAAAIALRGMAKDGKF
Vmp17 184 LFTATDAVRADDAAQAAAAASIGAVTGADILQAIVDSKENPVANSTDGIEKATDAAETAVAPAKDNKVE....IKDGAKKDAVIAAGIALRAMAKNGTF
VisE1 296 AVKDGEKEKAEGAIKGAAESAVRKVLGAITGLIGDAVSSGLRKVGDSVKAASKETPPALNK 356
Vmp17 280 SIKNNE.DAAVTTINSAAASAVNKTLSTLIIAIRNTVDSGLKTINEALATVKGEDKSVEAT 339

FIG.3A

VR-I
VR-II
VR-III
VR-IV
VR-V
VR-VI
Vis1
Vis2
Vis3
Vis4
Vis5
Vis6
Vis7
Vis8
Vis9
Vis10
Vis11
Vis12
Vis13
Vis14
Vis15
Vis16

FIG.3B

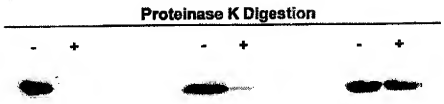

FIG.4A FIG.4B FIG.4C

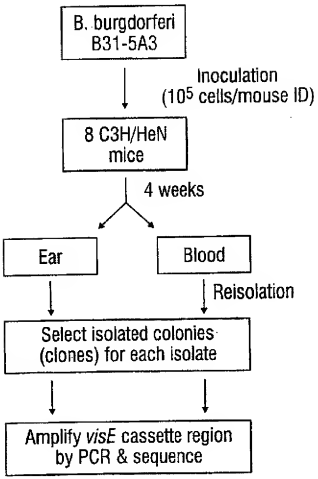
B. burgdorferi
B31-5A3
Inoculation
(10^5 cells/mouse ID)
8 C3H/HeN
mice
4 weeks
Ear
Blood
Reisolation
Select isolated colonies
(clones) for each isolate
Amplify visE cassette region
by PCR & sequence

FIG. 5A

VR-I
VR-II
VR-III
VR-IV
VR-V
VR-VI
V1sE1
K1e4A
K1b4A
M2b4A
M2e4A
M3e4A
M3b4A
M4e4A
M4b4A
M5e4A
M5b4A
M6b4A
M7b4A
M8e4A

FIG.5B

VR-I
VR-II
VR-III
VR-IV
VR-V
VR-VI
Vlse1
Mle4A
Mle4B
Mle4C
Mle4D
Mle4E

FIG.5C

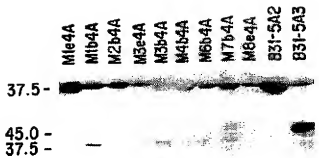

FIG. 6A

FIG. 6B

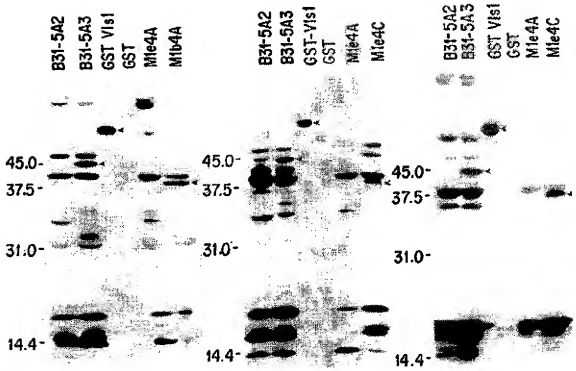

FIG. 6C FIG. 6D FIG. 6E

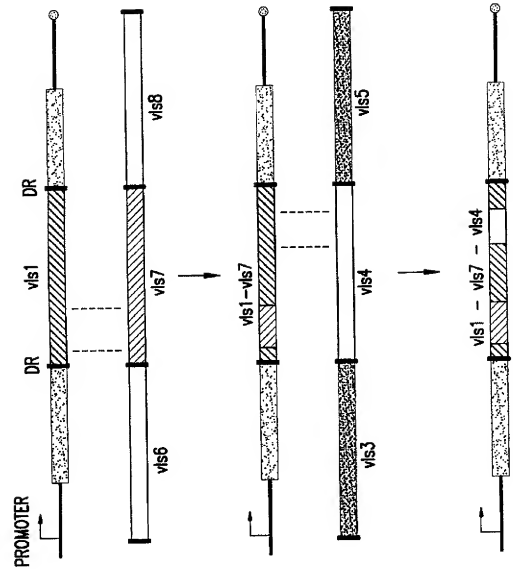
PROMOTER
DR
vls1
DR
vls6
vls7
vls8
vls1–vls7
vls3
vls4
vls5
vls1 – vls7 – vls4

FIG.7

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/02952

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :Please See Extra Sheet.
US CL :435/69.1, 71.1, 320.1; 536/23.1, 23.7
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/69.1, 71.1, 320.1; 536/23/1. 23.7

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

STIC

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, CAplus, MEDLINE, BIOSIS, EMBASE, SCISEARCH, WPIDS, DIALOG

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X ----- Y	RESTREPO et al. Subtelomeric Expression Regions of Borrelia hermsii Linear Plasmids are Highly Polymorphic. Molecular Microbiology. 1992, Vol.6, No.22, pages 3299-3311,see entire document.	1, 5 --------- 3, 27, 28
X	KITTEN et al. Intragenic Recombination and a Chimeric Outer Membrane Protein In the Relapsing Fever Agent Borrelia hermsii.Journal of Bacteriology, May 1993, Vol.175, No.9, pages 2516-2522.	1
T,E	ZHANG et al. Antigenic Variation in Lyme Disease Borreliae by Promiscuous Recombination of VMP-like Sequence Cassettes.CELL, 18 April 1997, Vol 89, pages 275-285, especially entire document.	1-12, 27, 28

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
02 JUNE 1997	25 JUL 1997

Name and mailing address of the ISA/US	Authorized officer
Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	VERLENE RYAN
Facsimile No. (703) 305-3230	Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/02952

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P	US 5,571,718 A (DUNN et al) 05 November 1996.	1, 5, 9-12
Y		7,8

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/02952

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-12, 27 and 28

Remark on Protest
☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/02952

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (6):

C12P 21/06, 21/04; C12N 15/00, 15/09, 15/63, 15/70, 15/74; C07H 21/02, 21/04

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claim(s)1-12, 27 and 28, sharing the inventive concept of a DNA, host cell, and method of using the DNA segment.

Group II, claim(s) 14-19, 22, 25, 26, 29, and 30, sharing the inventive concept of a polypeptide, and a method of immunizing by administering the polypeptide.

Group III, claim(s) 20, 21, 23, and 24, sharing the inventive concept of the antibody, and a method of detecting the protein.

Group IV, claim 13, sharing the inventive concept of a method for diagnosing Lyme disease comprising identifying a Vmp-like nucleic acid segment.

The inventions listed as Groups I-IV do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: In order for unity of invention to be present the claims must be linked by the same technical feature. A special technical feature is defined by PCT Rule 13.2 as a contribution over the prior art. The shared inventive concept of Group I is taught in Dunn et al (1992) and thus unity of invention does not exist with the corresponding special technical feature of Group II. Groups III and IV are drawn to distinct methods which also do not share the same inventive concept. PCT Rule 13 does not provide for a plurality of independent methods or products.